



GRAND CHALLENGES EMERGING PERSPECTIVES for Embedded Processing

Dennis Healy

Microsystems Technology Office

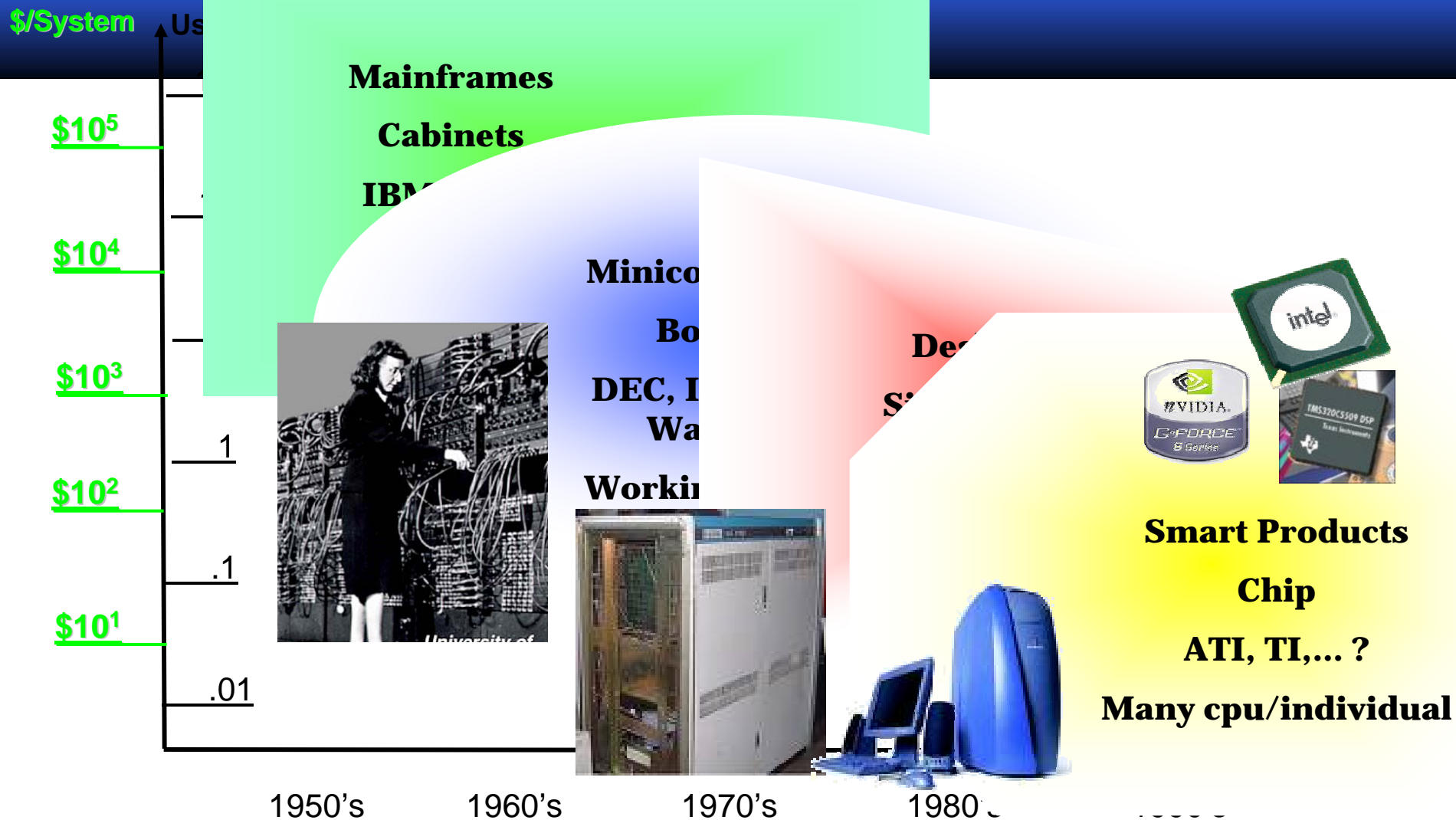
Defense Advanced Research Projects Agency

Dennis.Healy@darpa.mil

March 6, 2007

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 06 MAR 2007		2. REPORT TYPE N/A		3. DATES COVERED -	
4. TITLE AND SUBTITLE Grand Challenges Emerging Perspectives For Embedded Processing				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) DARPA				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited					
13. SUPPLEMENTARY NOTES DARPA Microsystems Technology Symposium held in San Jose, California on March 5-7, 2007. Presentations, The original document contains color images.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 46	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

The Rise of Blue Collar Processing



In DARPA's lifetime the user/cpu ratio has inverted

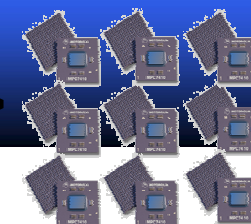


DARPA Director Lauds Winning Embedded Processing Solution to DARPA Grand Challenge



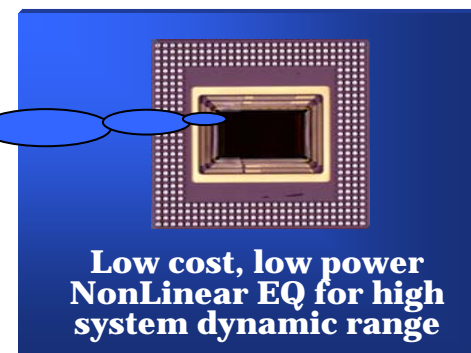
Photo courtesy of DARPA

- GC1: Embedded DSP Algorithms Mapping to non-traditional commodity architectures

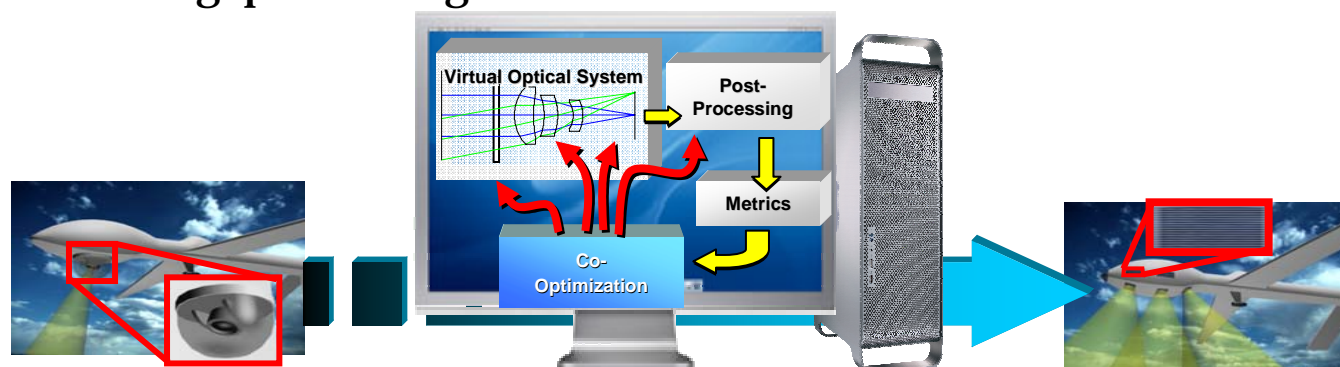


- GC2: Co-Design of Algorithms and Architecture
Towards Commodity Nonlinear DSP

$$y(n) = \sum_{p=0}^P \sum_{n_1=0}^N \cdots \sum_{n_p=0}^N H_p(n_1, \dots, n_p) x(n-n_1) \cdots x(n-n_p)$$



- GC3: Integrated Sensing with Processing
 - System level Co-design: Sensor/Processor/Algorithms
 - “compressed” sensing/processing





PROCESSING GRAND CHALLENGE 1

EO/IR Space Time Adaptive Processing (STAP)



Current airborne Advanced EO/IR System

1 m resol'n over 1km²

0.25 Mpixel @ 10 Hz

IR-STAP track before detect

Small targets in clutter



Figure 6-7. P2J128J CE Daughtercard (Heat Sinks Removed)

- real-time 4 PPC 7410 500 MHz processors, close to 4.93 GFLOPS
- 512 X 512 IR images at 2 Hz rate, 36 target velocity hypotheses

To move beyond TIVO in the sky:

Need TeraFlops Scale Embedded Processor

25,600 windows

5 x 5 x 5 matched filter

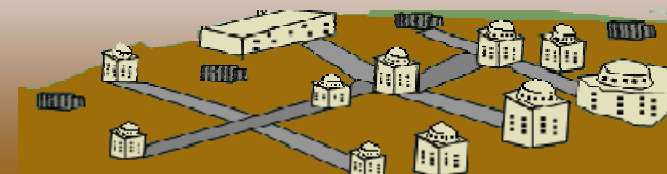
10 target models

10 second update

10,000 targets

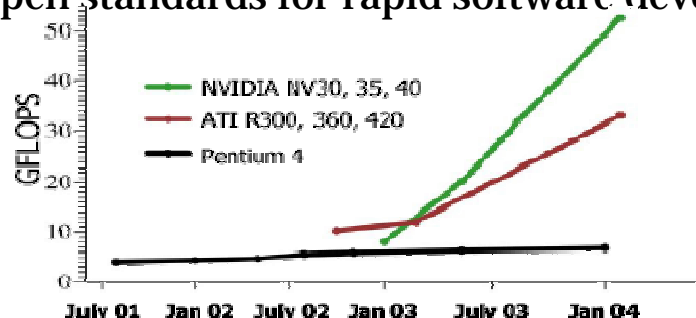
Upcoming Systems

- 100 Mpixels @ 2 Hz



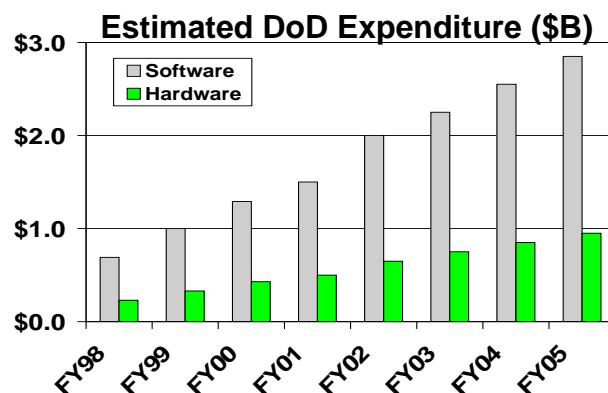
Gaming hardware/software explosion

- *Optimized for complex geometric forward problems*
- Big Market – fastest growing commercial I.T. drives
 - Increased capability, programmability
 - Towards portable form factor, power
 - Open standards for rapid software development



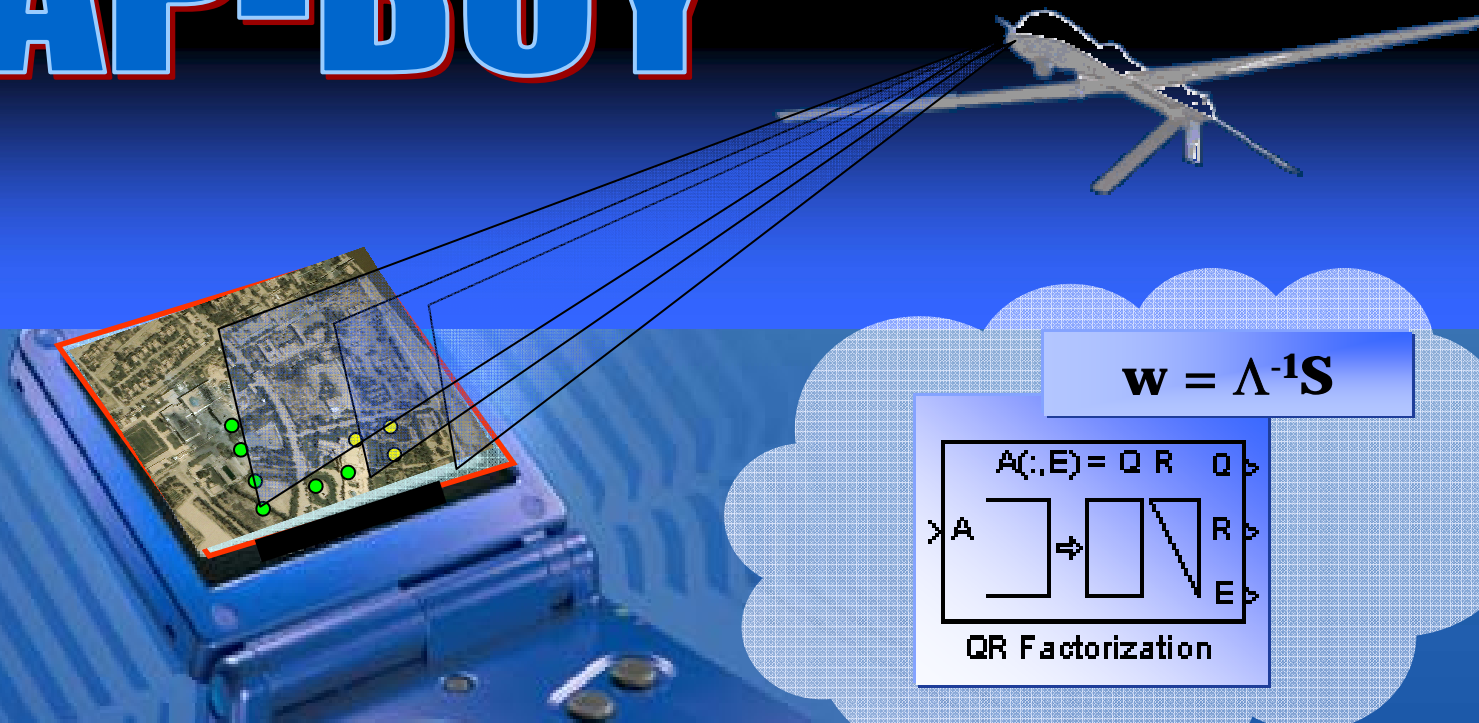
DoD Sensor processing challenges

- *Complex geometric inverse problems*
- Form Factor Limitations
- Curse of the small market
 - Hardware \$\$
 - Software \$\$\$\$

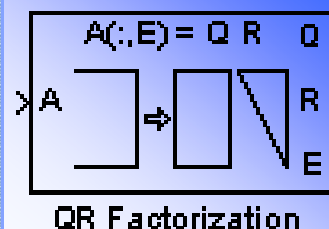


STAP-BOY

(STAP = Space-Time Adaptive Processing)



$$\mathbf{w} = \Lambda^{-1}\mathbf{S}$$



Low Cost Tensor Processor Architecture enables algorithms
from Gaming to Geometry-based Wave concepts



- Algorithm Mapping (for algebra) to Geometry-based DSPs
- 3D Pipeline Mesh Scalable Architecture
 - 1 COTS Graphics Processor Unit (GPU)
- Transition to Low Power Handheld Computing for Disruptive Chip
 - > 50 GFLOPs
 - > 10 AltiVec Power PC DSPs
 - 2.5 GFLOPs



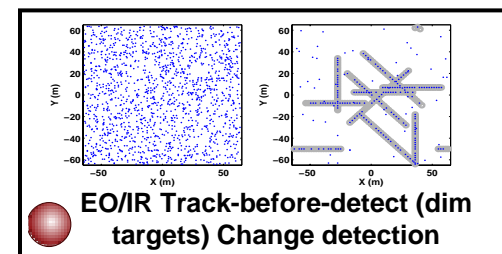
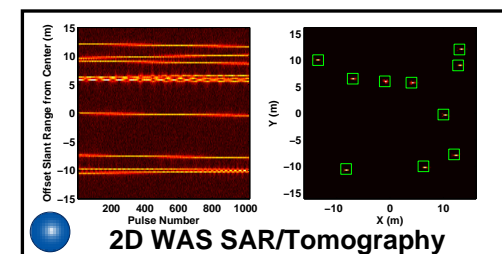
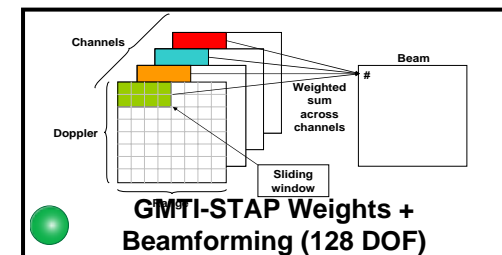
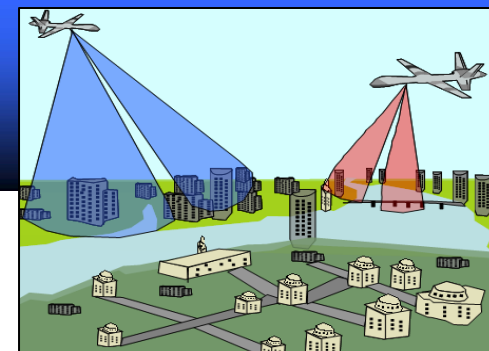
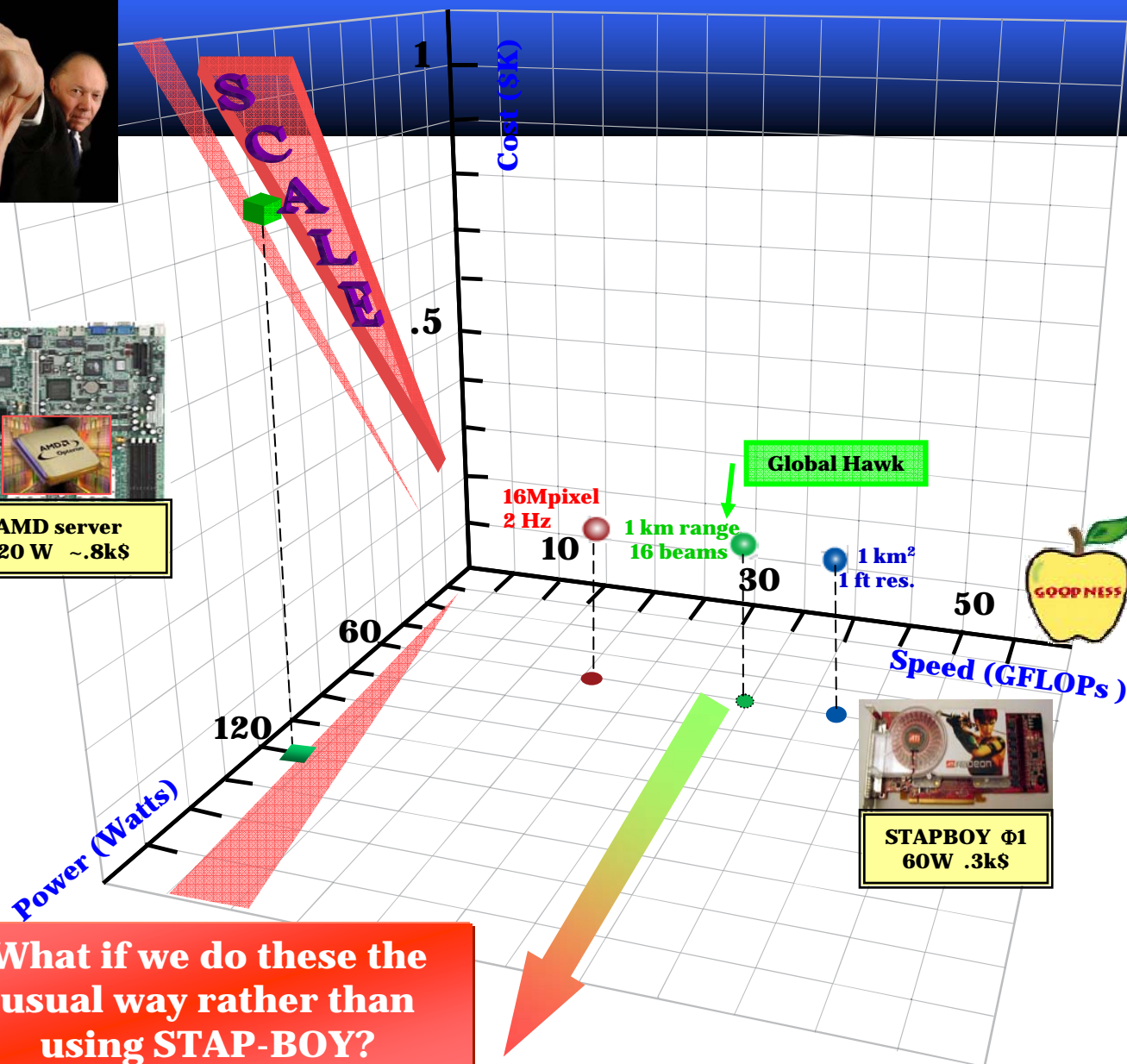
Phase I in Perspective: *STAP-BOY* in “The Embedded Sensor Processor Trade Space”



Badness



AMD server
120 W ~.8k\$



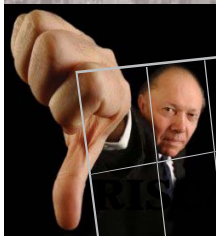
What if we do these the usual way rather than using STAP-BOY?



Phase I in Perspective: *STAP-BOY* in “The Embedded Sensor Processor Trade Space”



Badness



DSP

SCALE
SCALE
SCALE



AMD server
120 W ~.8k\$

STAPBOY $\Phi 1$
60W .3k\$

STAPBOY $\Phi 1$
60W .3k\$

Global Hawk
MPRTIP

ixel
1 km range
16 beams

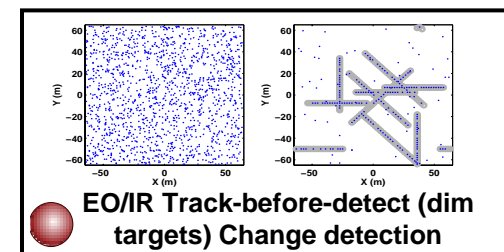
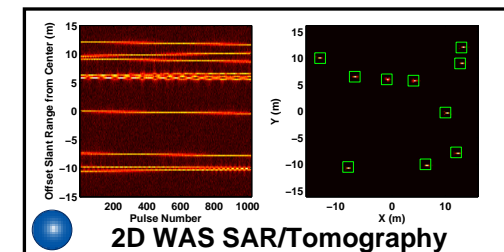
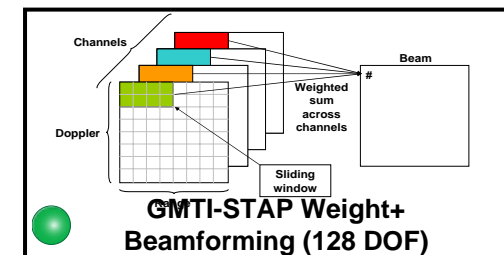
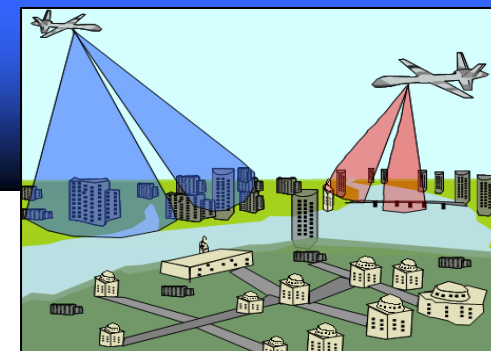
1 km²
1 ft res.



Speed (GFLOPs)

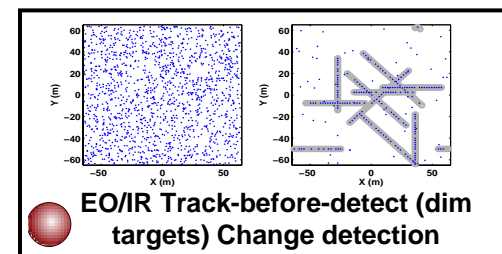
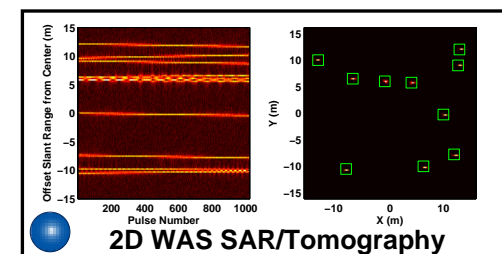
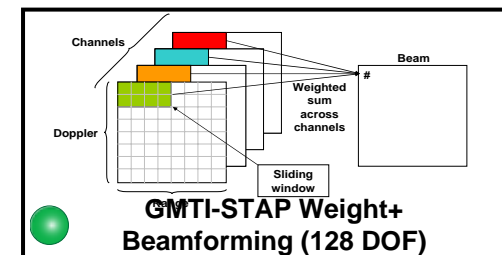
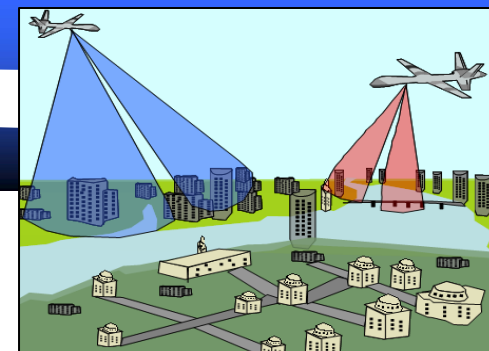
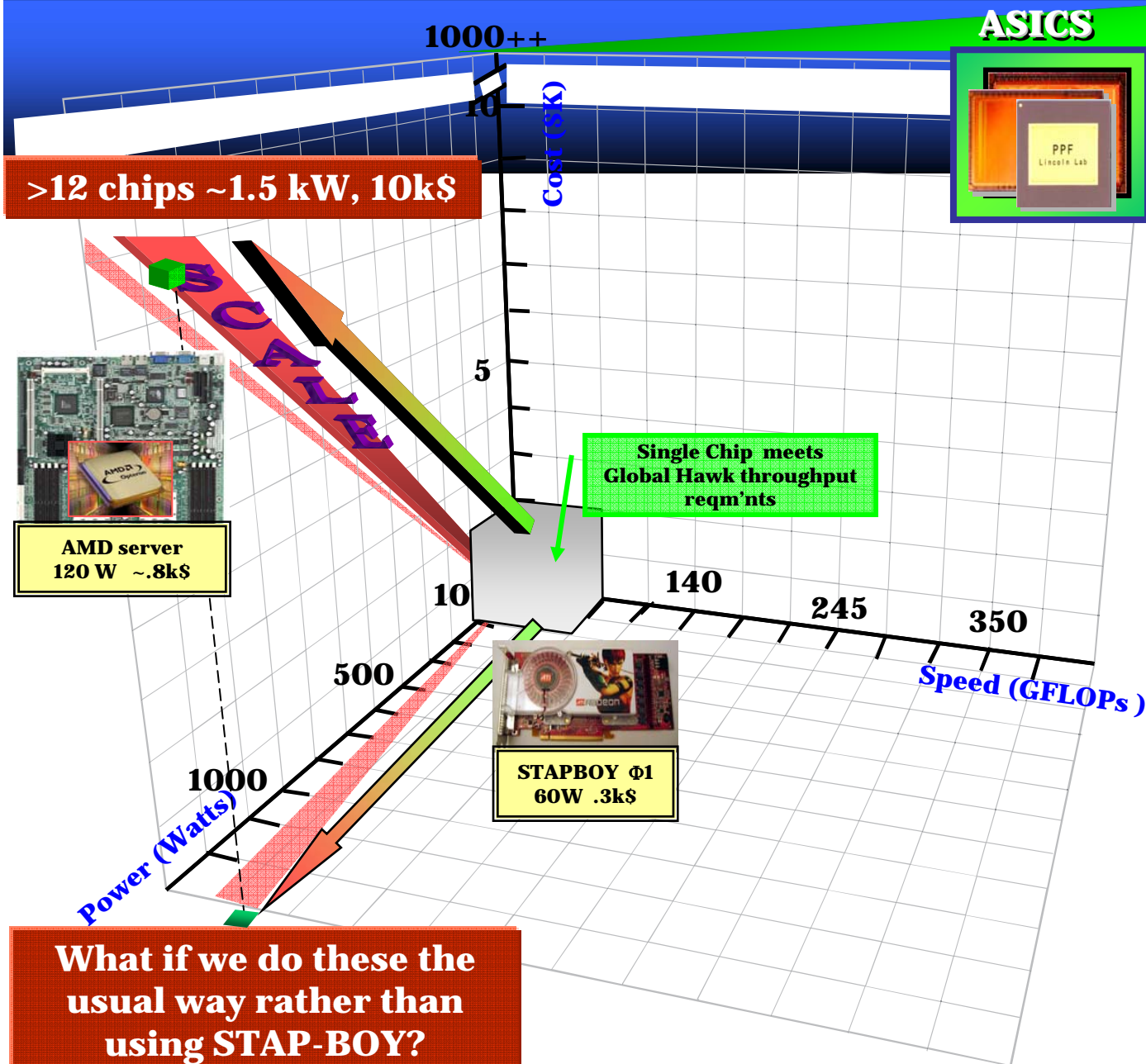
Power (Watts)

Cost (\$K)

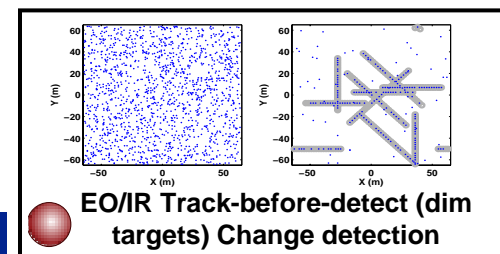
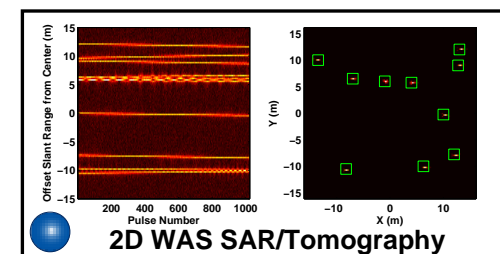
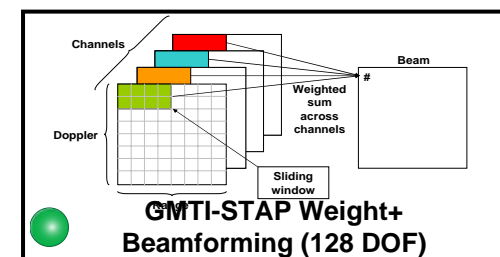
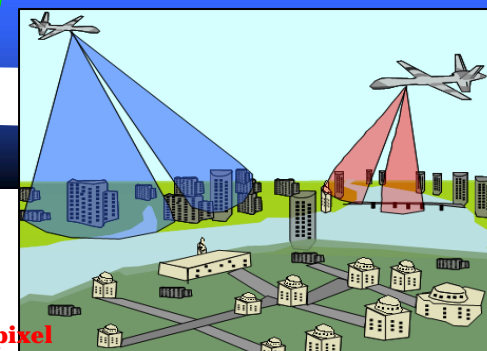
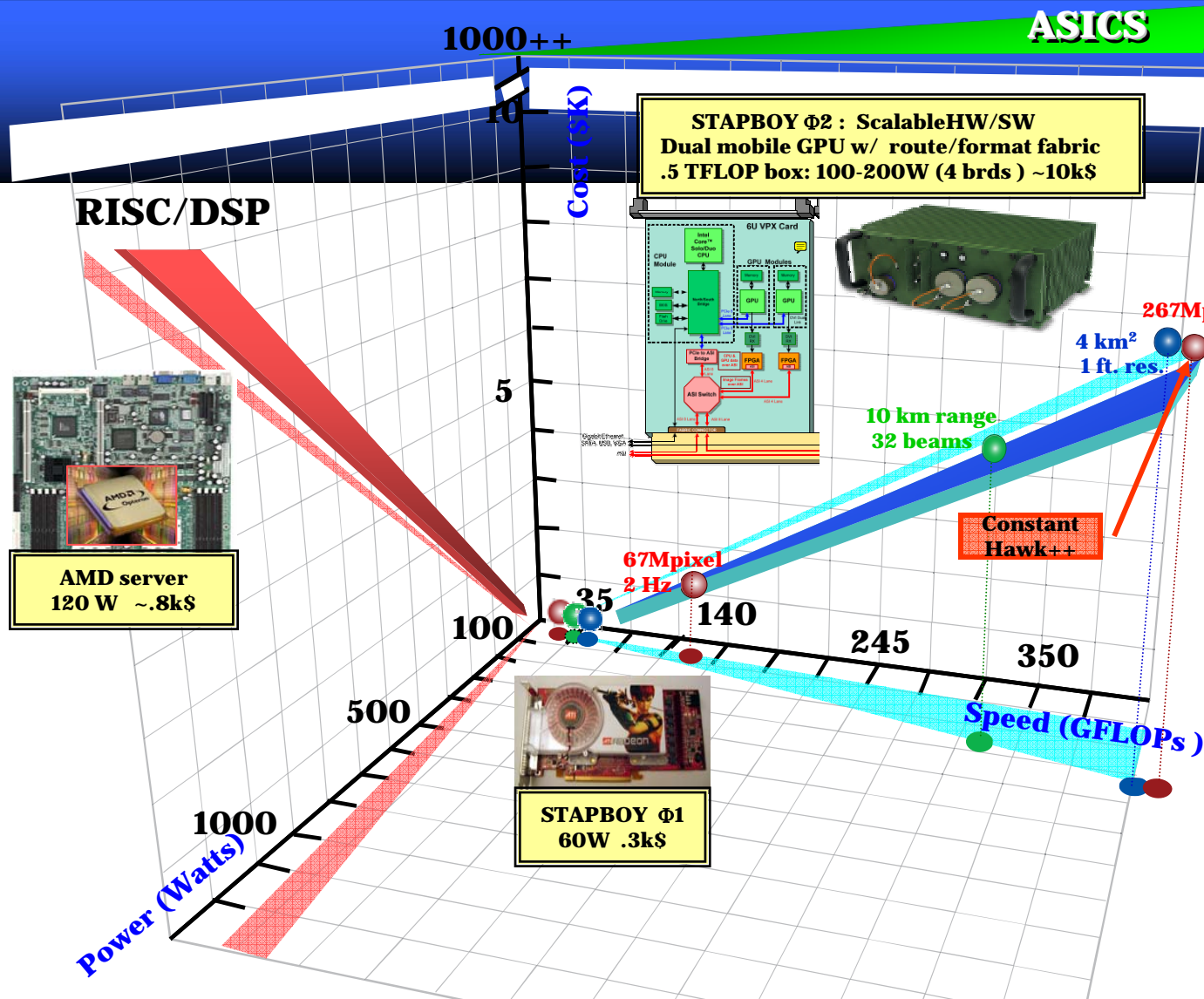


What if we do these the usual way rather than using STAP-BOY?

Phase I in Perspective: STAP-BOY in "The Embedded Sensor Processor Trade Space"



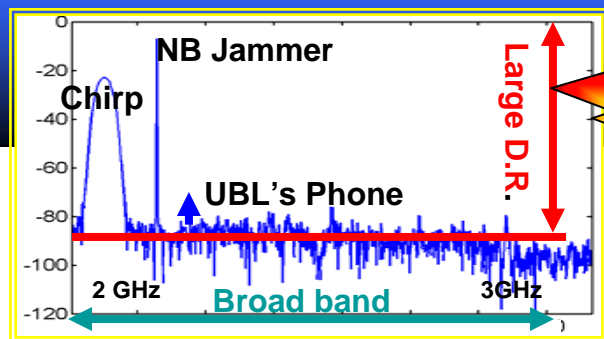
Phase 2: Can we scale the *STAP-BOY* low cost low power hardware and software to required system performance?



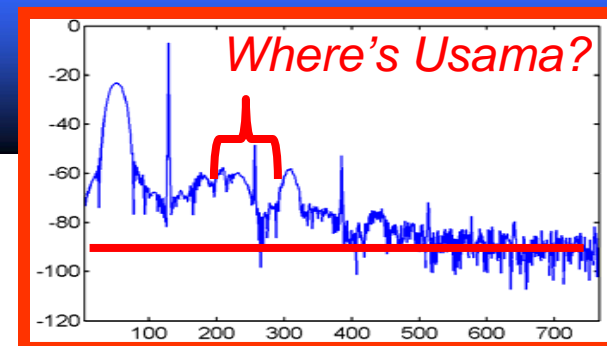
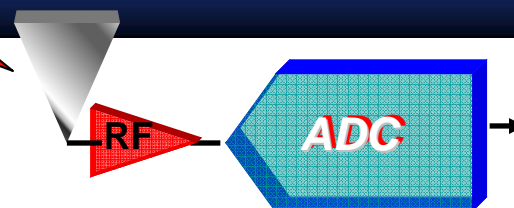
Next Challenge:
Advancing Sensor Resolution Drives Processing Well Beyond CPU/DSP and ASIC Envelopes for Cost & Comms Constrained Platforms

PROCESSING GRAND CHALLENGE 2

Getting a clear view through the RF front end



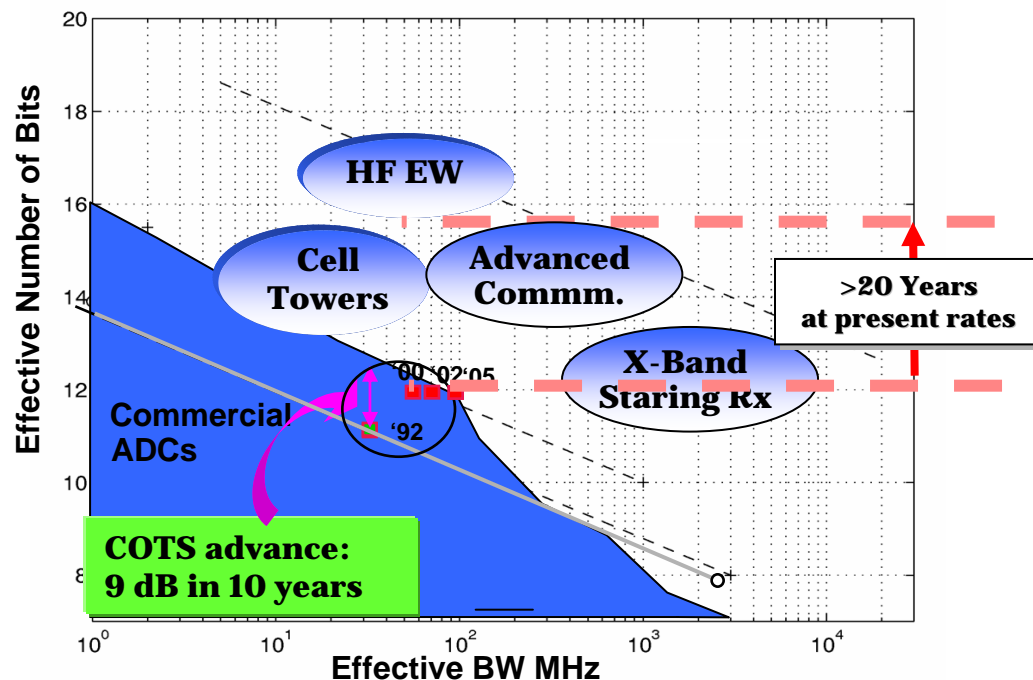
Complex RF environment



Distorted digital Representation

Rapid Proliferation of LPI: UWB, spread, hopped, coded...

Digital Receivers not keeping up



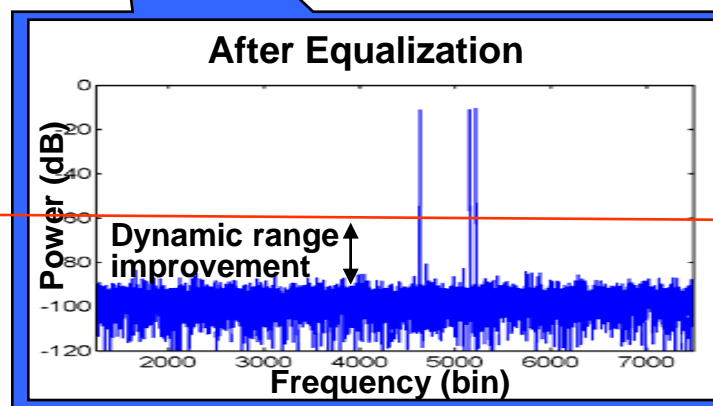
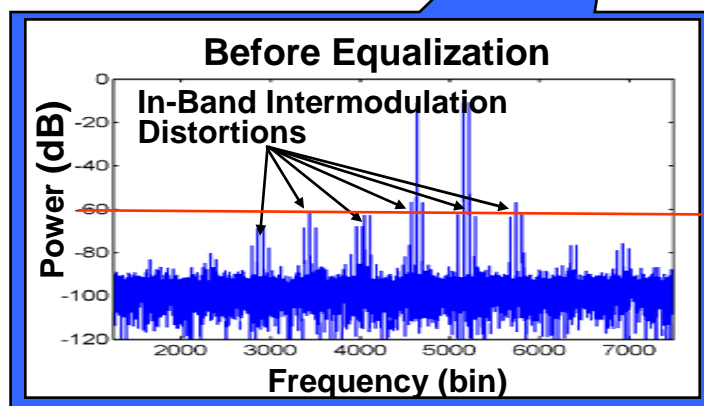
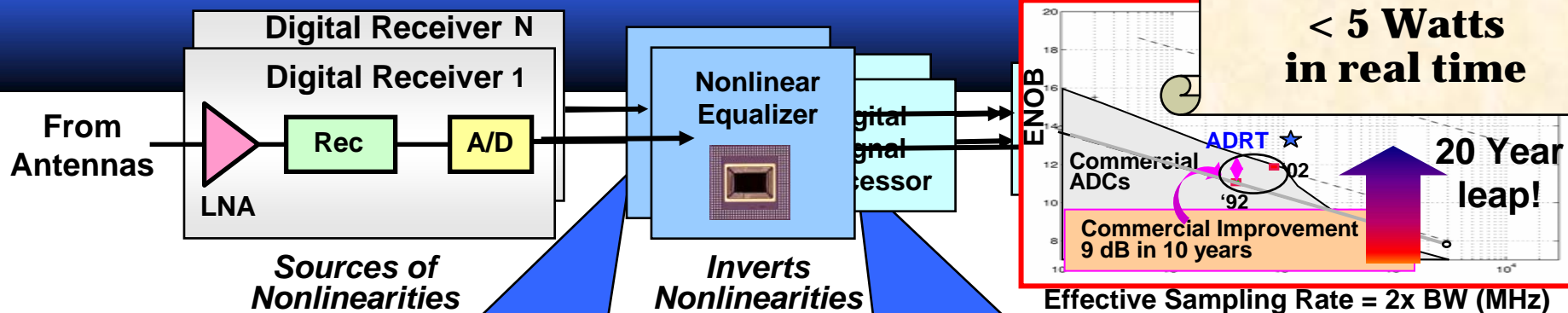
Attacking this problem through a revolution in nonlinear digital signal processing



Co-Designed Nonlinear Equalization (NLEQ) Advanced Signal Processing for Enhanced Dynamic Range

Program Goal
+ 25 dB IFDR
@500MHz IBW
< 5 Watts
in real time

Representative System

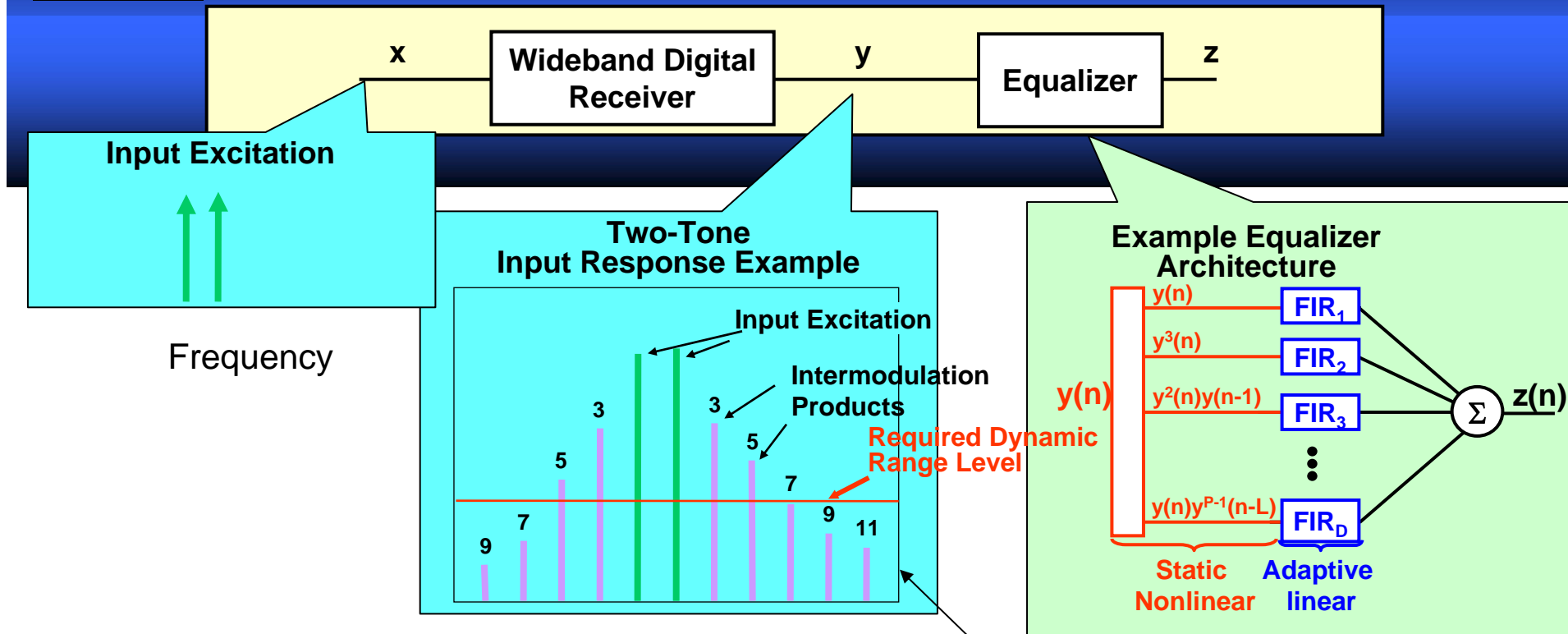


Objective: Overcome sensor device performance and cost limitations by **VLSI implementation of advanced nonlinear algorithms**

Key Issues: The Curse of Dimensionality

- Computational Burden of NLEQ
- Training Cost of NLEQ

Approach: **Compressed Representations. Algorithm/Hardware co-design**

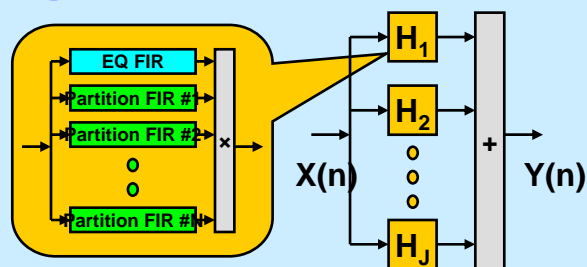


$$y(n) = \sum_{p=0}^P y_p(n) = \sum_{p=0}^P \sum_{n_1=0}^N \cdots \sum_{n_p=0}^N C_p(n_1, \dots, n_p) x(n - n_1) \cdots x(n - n_p)$$

- Generalization of Taylor series ($N=0$)
- Generalization of linear FIR filter ($P=1$)
- Naively, size scales as N^P for memory N and nonlinear order P
- Training NLEQ looks like a high-d optimization problem

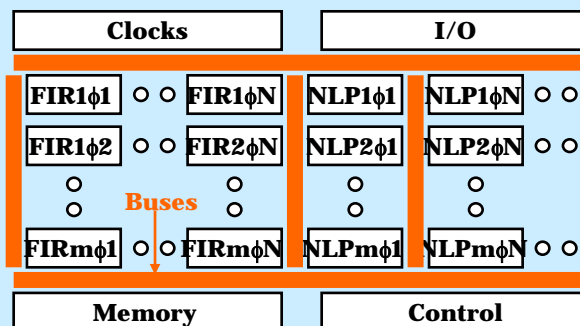
Co-Design of algorithm and hardware: Low-power VLSI NLEQ Implementation

NonLinear Equalizer Algorithm/Architecture*



PHoCS Compressed Nonlinear representation
Distributed Polyphase Block-Floating-Point Residue Arithmetic Architecture
500 - 2000 OPS/Sample
25,000 – 100,000 Low-Power Multipliers
20,000 – 80,000 Accumulators

IC Architecture and Floor Plan*



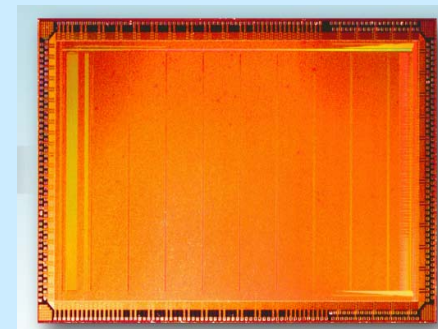
Systolic architecture minimizes high-speed comm. path lengths

High clock rate processing across entire die area

Two Input Ports, Two Output
750 MSPS Demultiplexed
1500 MSPS Full Rate

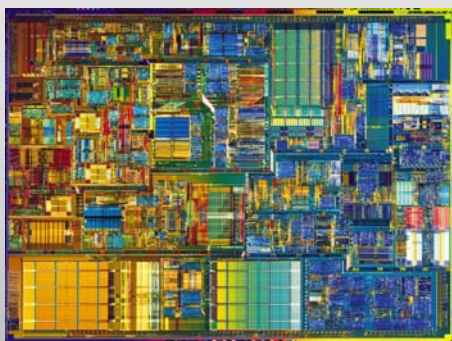
Control

Processor Die



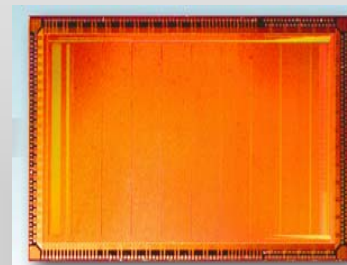
0.09 - 0.13 Micron CMOS
30M - 100M Devices
2 - 4 cm² Die Area
Low-Power Low-Vt Dynamic Logic
~1TOPS, <3 Watts
Low-power LVDS compatible I/O drivers

Pentium 4 die

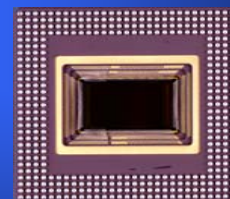
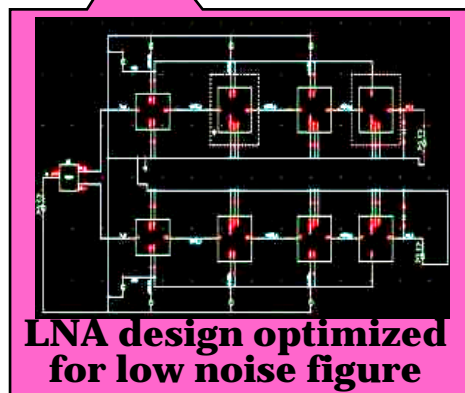
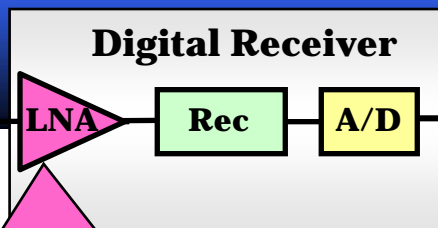


- 1300 dies required for 1 Teraops NLEQ
 - 150 KW total power
- Assumptions for each die
 - 3.8 GHz
 - 115 Watts
 - 10% code efficiency

Phase 2 NLEQ DSP Die

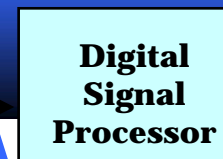


- 1 die required for 1 Teraops NLEQ
 - <5 W total power
- 50,000x power efficiency
- 1,300x computational throughput density
- 65,000,000x total figure of merit of computational density times power efficiency



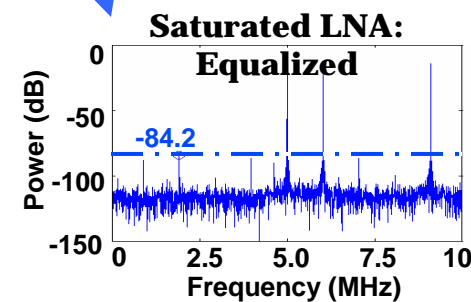
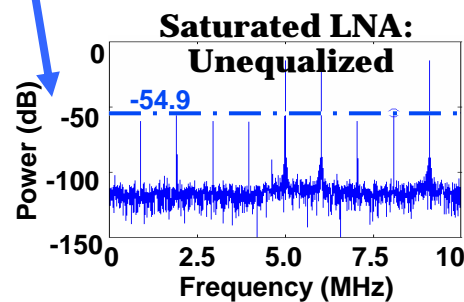
Low cost, low power digital NLEQ provides necessary high system dynamic range

$$y(n) = \sum_{p=0}^P \sum_{n_1=0}^N \cdots \sum_{n_p=0}^N H_p(n_1, \dots, n_p) x(n-n_1) \cdots x(n-n_p)$$

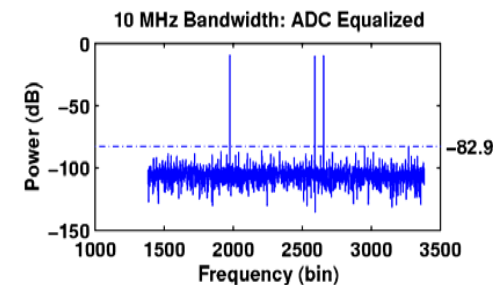
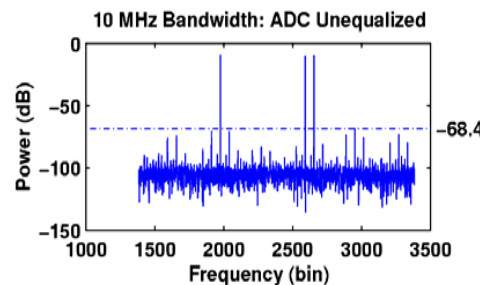


Current circuit designs can be costly when optimizing both low noise *and* high dynamic range

S-Band Receiver and ADC (10-MHz Bandwidth)



ADC Only (10-MHz Bandwidth)

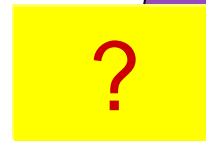
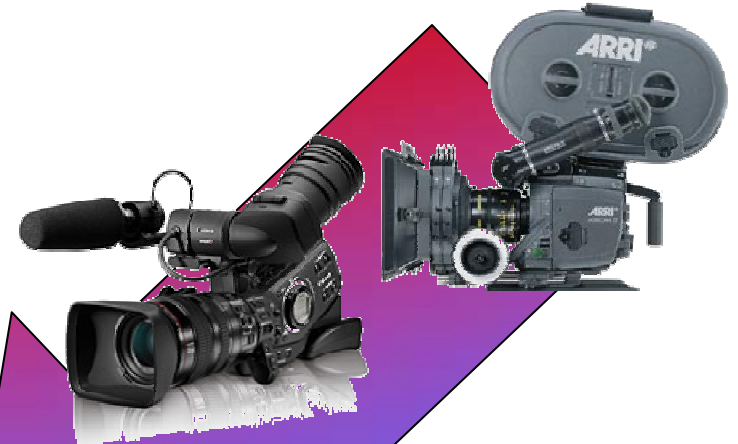
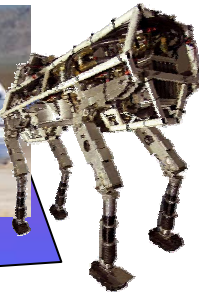


Embedded Non Linear Signal Processing simplifies circuit designs if used early in the design process

Brains and bodies evolved together

- Embedded processing is optimized for the needs of the system, say a sensor system like a camera. Why can't the camera be optimized for the processing?

*"Cameras will also change form. Today, they are basically **film cameras without the film**, which makes about as much sense as automobiles circa 1910, which were **horse-drawn carriages without the horse**. A car owner of that time would be pretty shocked by what's in a showroom now. Camera stores of the future will surprise us just as much."* -Nathan Myhrvold, former CTO of Microsoft, co-founder of Intellectual Ventures, NY Times, 5 June 2006



- Structural & functional impact of co-design over the *entire* system?

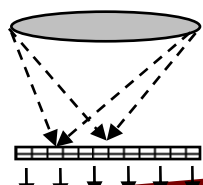
Co-Design opens a new design space between Camera and Computed Imaging



**Optical Field
Processing**

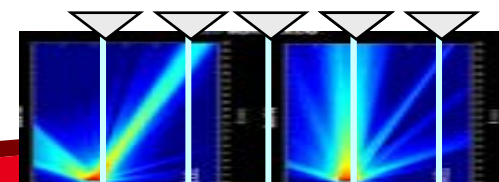


**RF Pixel
Sampling**



**Digital Pixel
Processing**

Image



Digital Beamformer RF Array

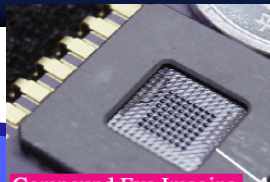
**Digital Field
Processing**

Image

“Load Balancing” between Photon Processing and Bit Processing



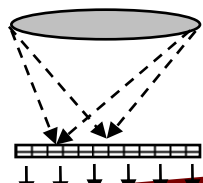
Optical Field Processing



Optical Regional Processing

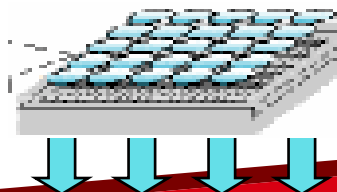


RF Pixel Sampling



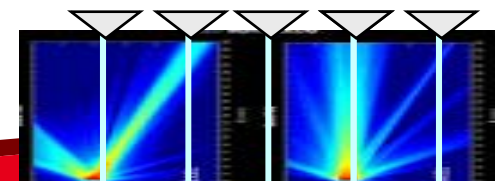
Digital Pixel Processing

Image



Digital Regional Processing

Image



Digital Beamformer RF Array

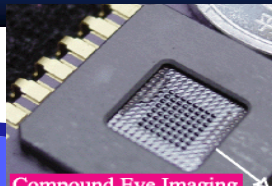
Digital Field Processing

Image

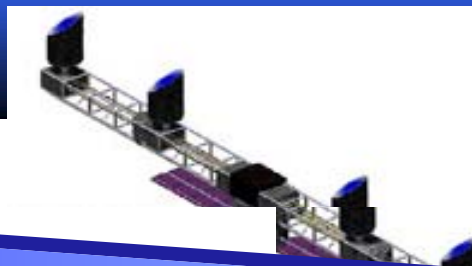
“Load Balancing” between Photon Processing and Bit Processing



Optical Field Processing



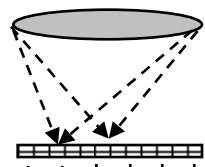
Optical Regional Processing



Optical Pixel Processing

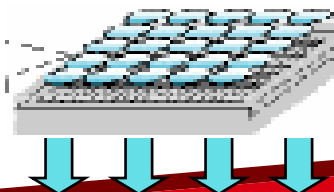


RF Pixel Sampling



Digital Pixel Processing

Image



Digital Regional Processing

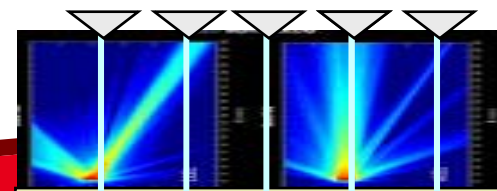
Image

Sampling Field Sensor



Digital Field Processing

Image



Digital Beamformer RF Array

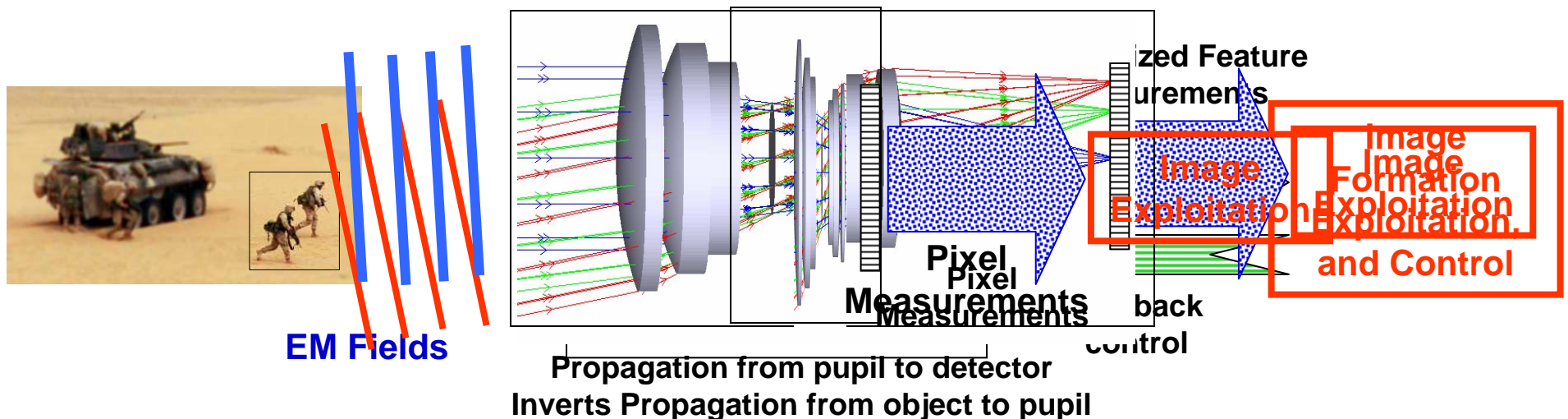
Digital Field Processing

Image

Exploitation

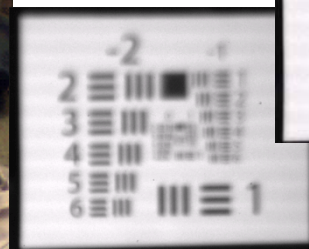
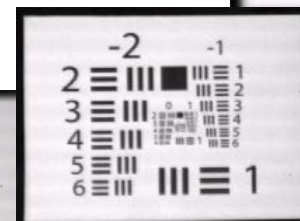
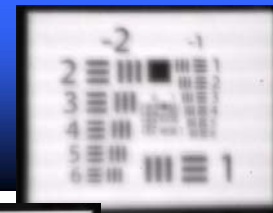


Radically transform form, fit and function of imaging sensors
 Integrated systems for transforming photons to information
 Simplified Manufacturing and Integration Process



APPROACH:

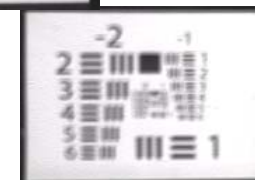
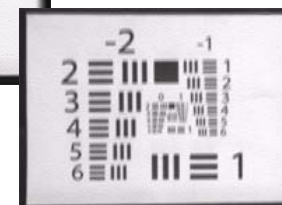
- Joint optimization of optics, sensors, post-processing algorithms
- Unconventional wavefront mapping (well conditioned encoding)
- Information-rich parameters for (compressive) measurement
- Integrated and simplified manufacturing approaches



2.2m

In focus (2.6 m)

3m



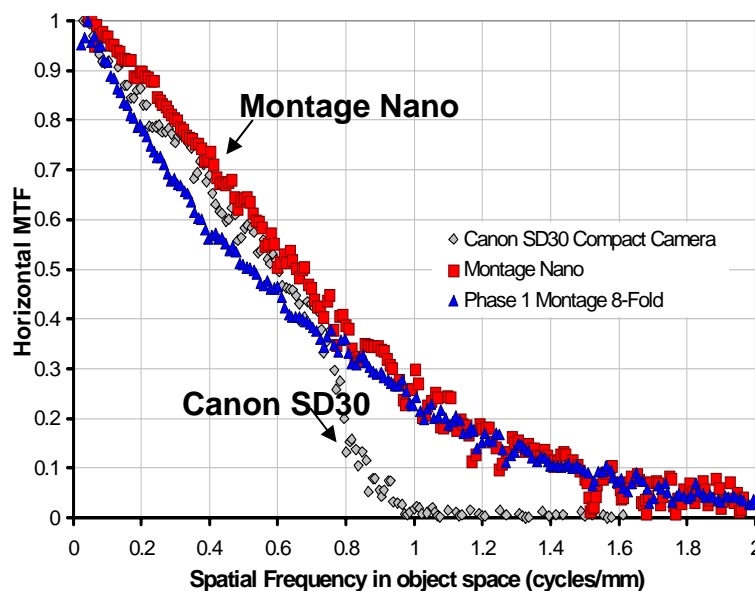
Montage Nano vs Canon SD30: Resolution Results

Canon Powershot SD30

5 Megapixel compact camera, 120g



	Canon SD30	MontageNano
Optical track	~30mm	5mm
Light collection	< 20mm ²	80mm ²
Resolution (2.2m range)	~30 lp/mm	100 lp/mm
Focal length	6.3 – 14.9mm	43mm
Optics	Multi-element refractor	Single element folded reflector

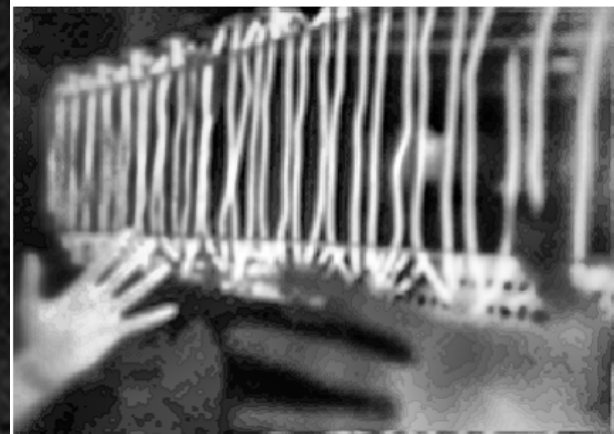
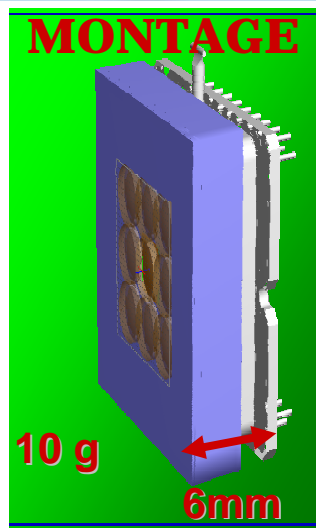
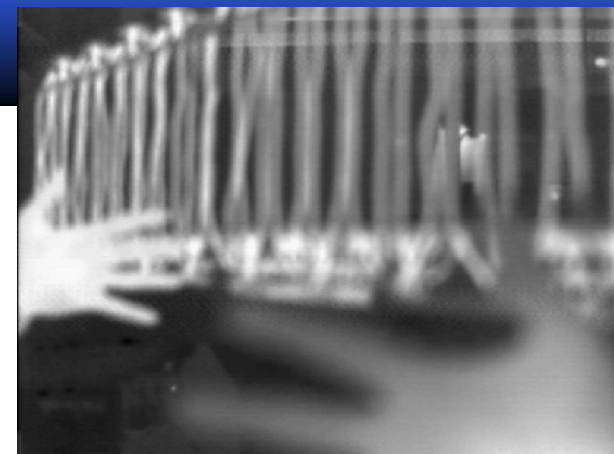
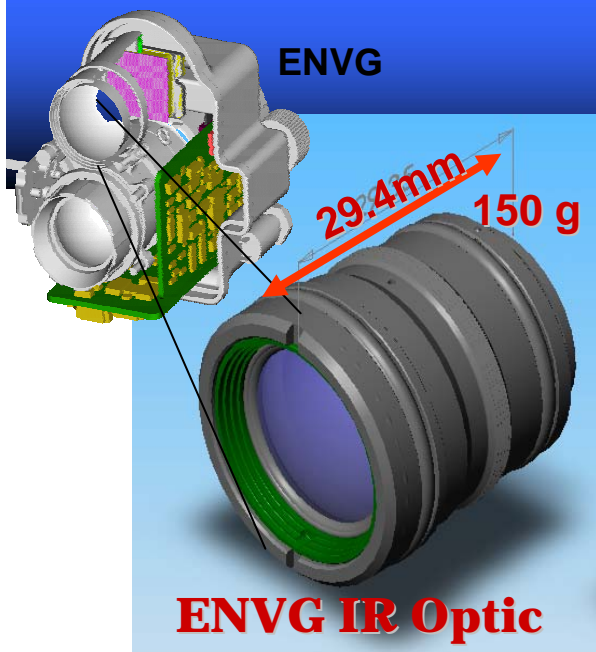


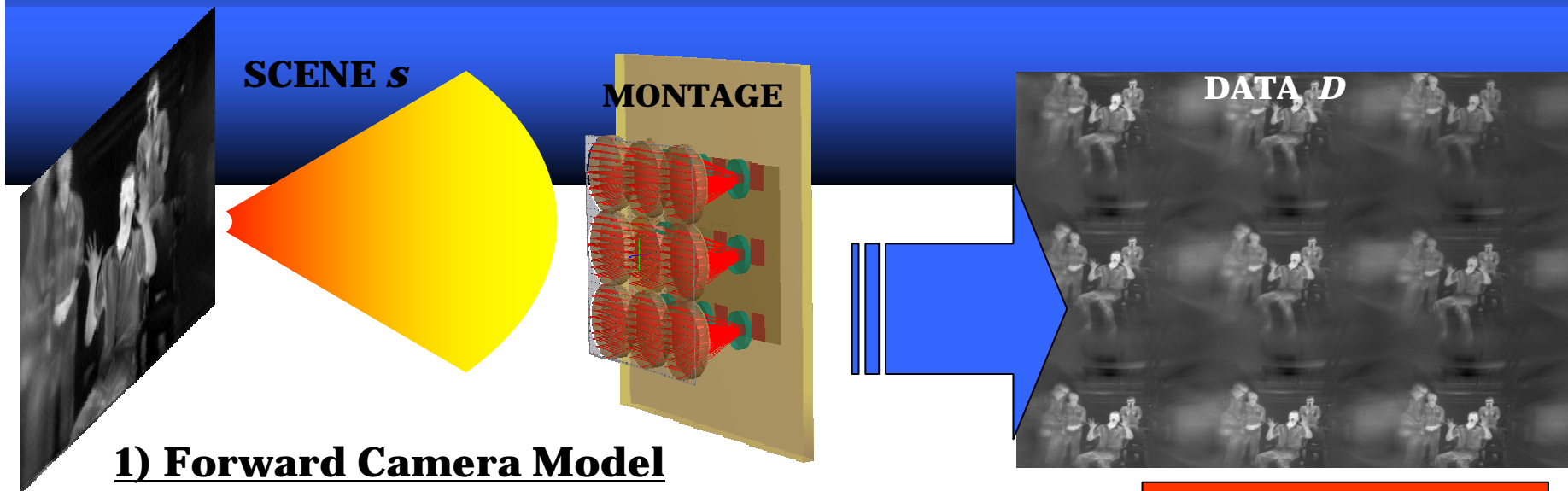
Montage MDO "Nano"

2 Megapixel, **ultra**-compact camera 30g
(< 15 g w/o metal housing)



Lenslet Array LWIR camera vs. SOA





1) Forward Camera Model

Predicts camera data D produced by a given scene

2) Solve Imaging Inverse Problem

Given data, estimates scene that produced it

$$\hat{s} = \arg \min_{s \in S} \left\{ L(s; D) + P(s) \right\}$$

Discrepancy between observed Data D & Data the Camera Model predicts from s

Penalty for unlikely or unphysical choices of s

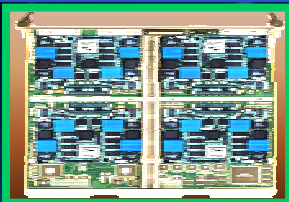
Sparsity Regularized Inversion



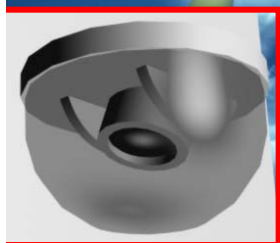
Reconstructed Image

Innovations: COMPRESSIVE ENCODING
CO-DESIGN OF CAMERA & INVERSE ALGORITHM
SPARSITY AS CONSTRAINT

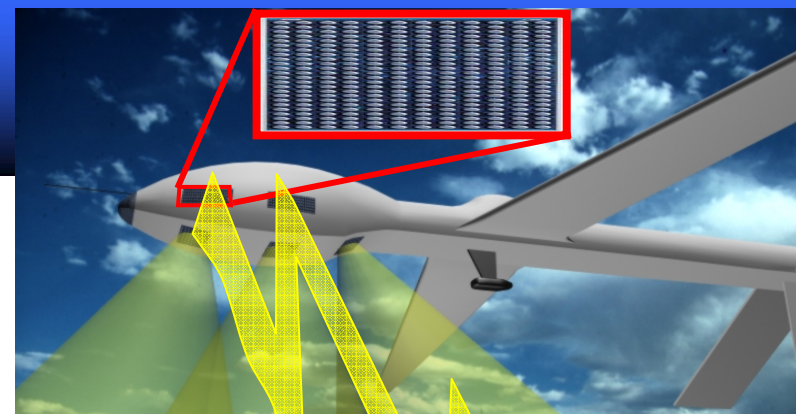
Exploiting Sparsity in Compressed Sensing/Processing



Compress: Compute hard, throw away 80% of the Megapixels for Downlink through tiny pipe



High Resolution Image :
Big Optics, Megapixel Data



Compressed Camera *directly measures*
Compressed representation for downlink

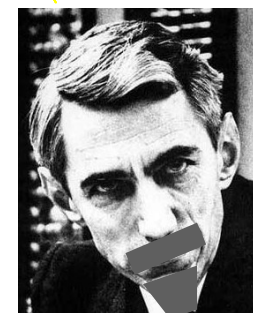
Uncomfortable Observation:

Compressibility means most of the data in the megapixels is useless

BUT

The gold is hidden in a broad spatial bandwidth

Shannon: must measure the megapixels to capture the hidden gold
then sieve it out with digital processing. **Pessimistic!**



Compressive Sensing (Analog-to-Information):

Nonadaptive Measurement at 20% of Megapixel rate gets the gold
Simpler Sensor. No Digital Compression on the platform.



Power to the Processors (right on!)

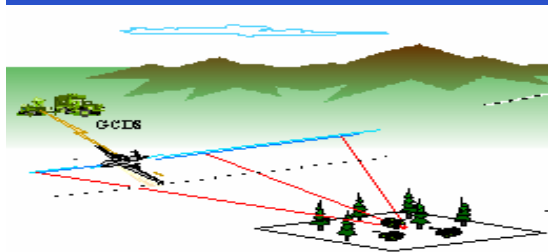


Photo courtesy of DARPA

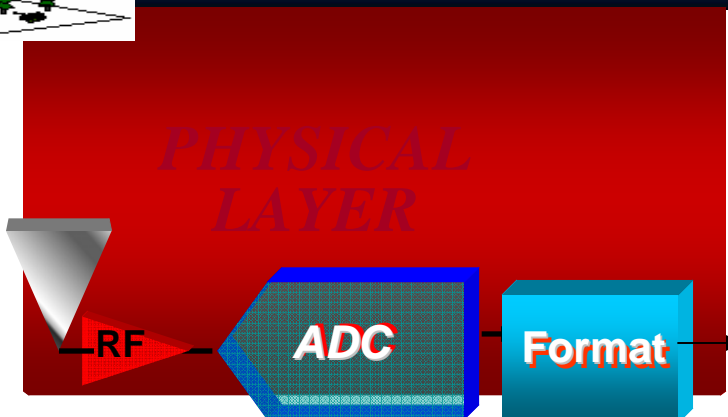
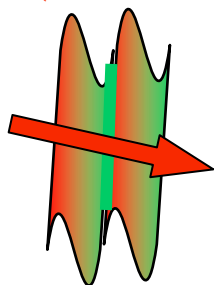


Fin

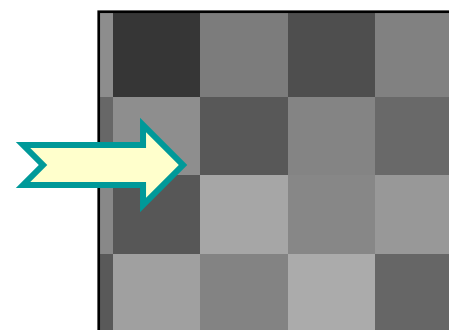




Physical Field
(continuum)

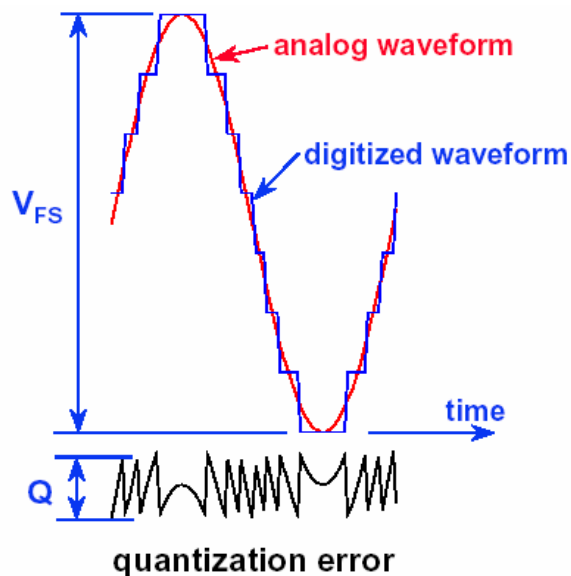


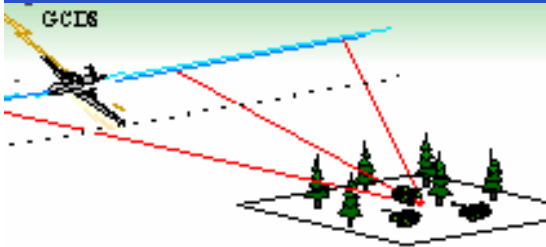
Digital Representation
(Formatted)



Raw sensor data:

*Huge Volume
low information/sample
Curse of Dimensionality*



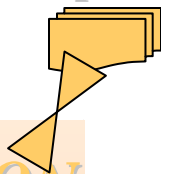
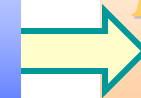
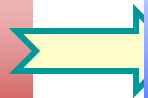
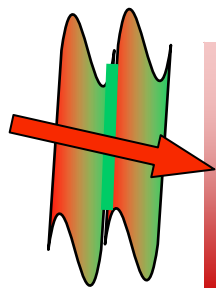


Physical Field
(continuum)

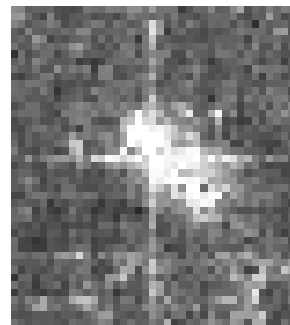
Digital Representation
(finite precision finite dimensional)

Transformed
Digital Representation

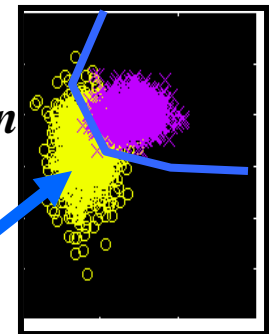
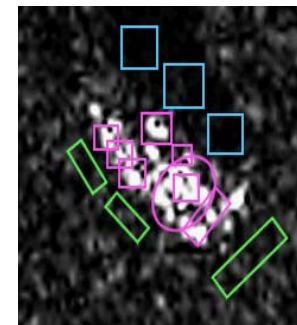
Symbolic Output



Raw sensor data:
low information/dimension

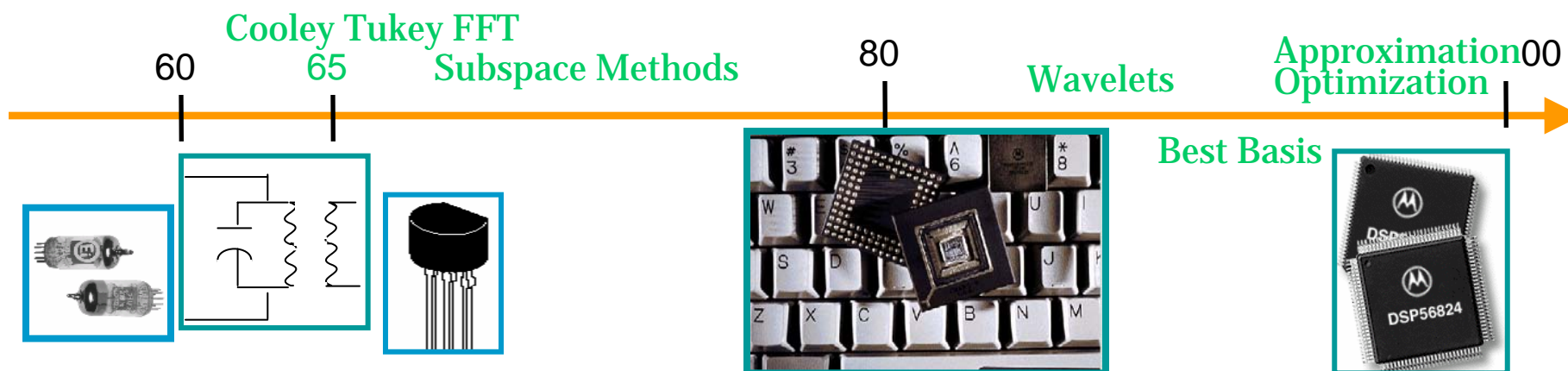
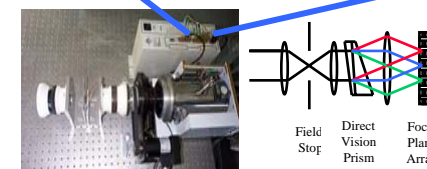
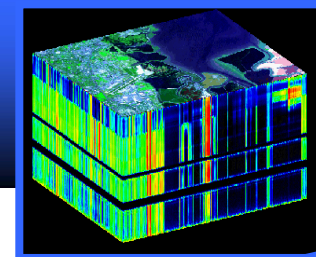
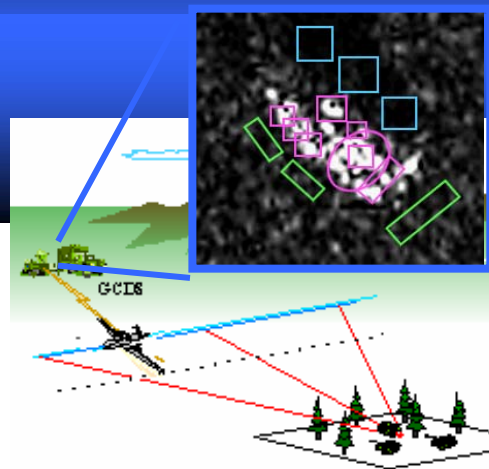
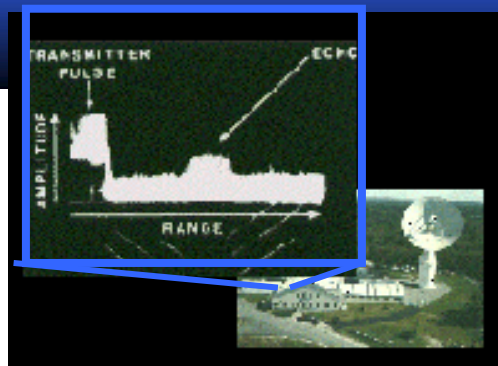


Enhanced feature data:
more information/dimension



Sensor: Work hard in the physical layer to measure everything at the highest fidelity

Processor: Work hard to throw away most of that data to get at the “good stuff”



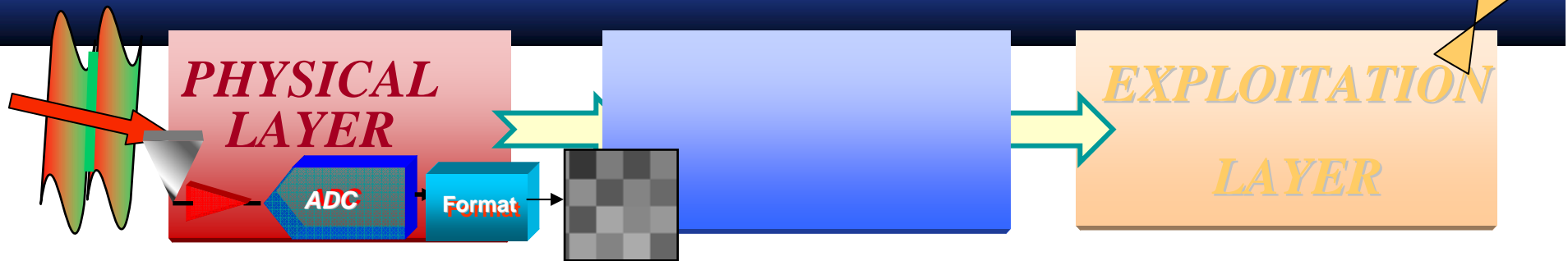
- *Increased complexity, dimensionality of sensing and exploitation tasks*
- *Decoupling of Digital and Analog subsystems, independently designed and optimized*
- *Steady Moore's Law growth of digital throughput (integration growth path)*
- *Equal impact of fast algorithms (representations), but sporadic, unpredictable*

Physical Field
(continuum)

Digital Representation
(finite precision finite dimensional)

Transformed
Digital Representation

Symbolic Output



Keep trying to measure/digitize everything at highest possible resolution?

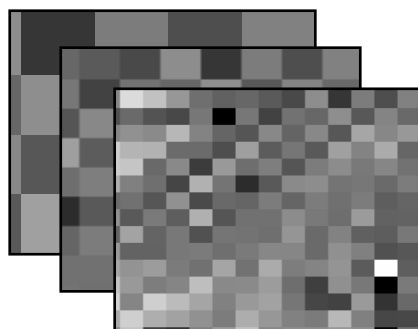
- A No-brainer
- No-brainer! Maybe not always true!
 - Slow, expensive growth rate of resolution in A/D's
 - Curse of dimensionality: shouldn't do it even if you could!
- What might be better?
 - Dedicate Front end resources to measuring more of the “good stuff”
 - Sensing “features” at front end could reduce load on A/D and subsequent digital processing and exploitation.
 - If done properly, overall system performance can be improved even with the lower requirements on subsequent digital processing
- **Problem: the “smarts” for finding the “good stuff” are behind the ADC!**

Today's Sensor Systems are typically feed-forward networks for transforming information in specialized stages.

Dimension reduction is mainly applied at the back of the bus



*Raw sensor data:
low information/dimension*

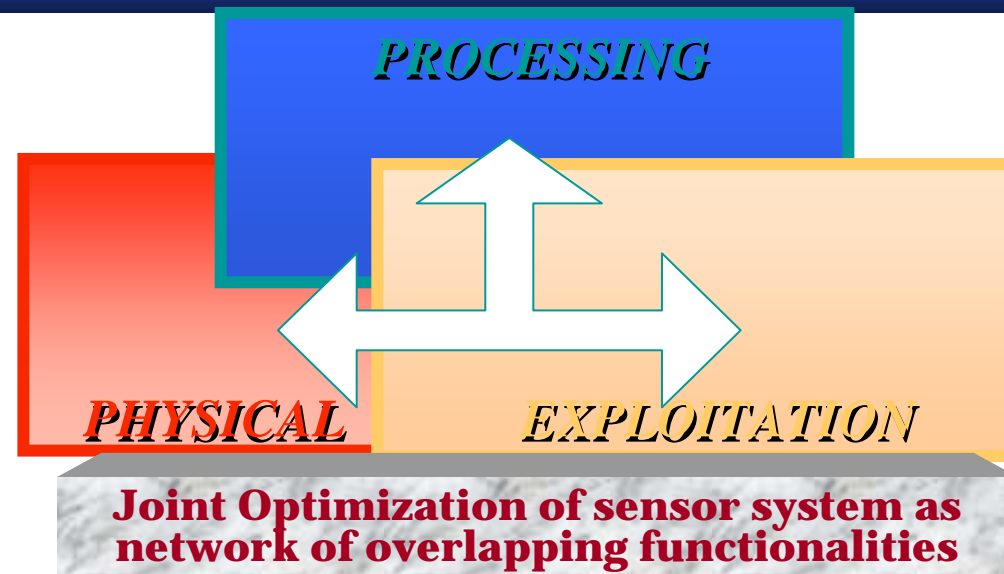


*DSP produces data reps with
higher information/dimension*



*Fast approximate
solution of a high
dimensional non
linear optimization
problem*

Meeting the Challenges: *Integrated Optimization of the System*



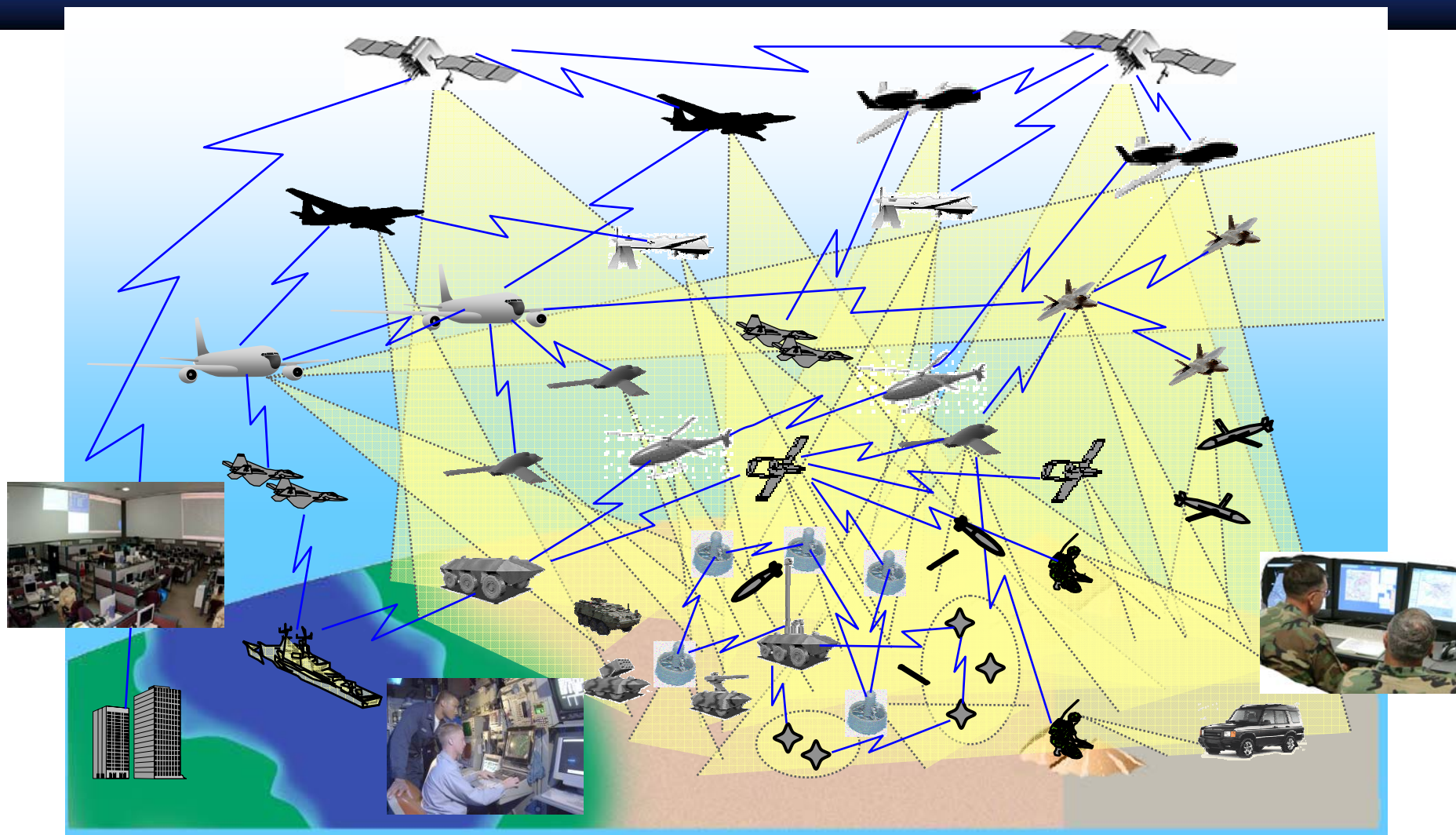
i

*Components will have overlapping and dynamically reconfigurable roles and full network connectivity. (**LOAD BALANCING**)*

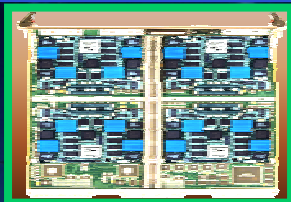
*Make “customized” measurements at physical layer under real-time feedback control from the exploitation and processing system. **“20 Questions”***

*Manage/Prioritize **data** stream to affordable levels without discarding needed **information***

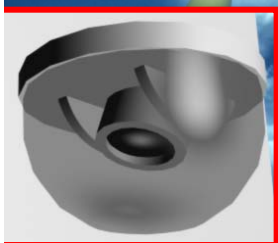
If that's not bad enough, we are building an internet of the damn things....



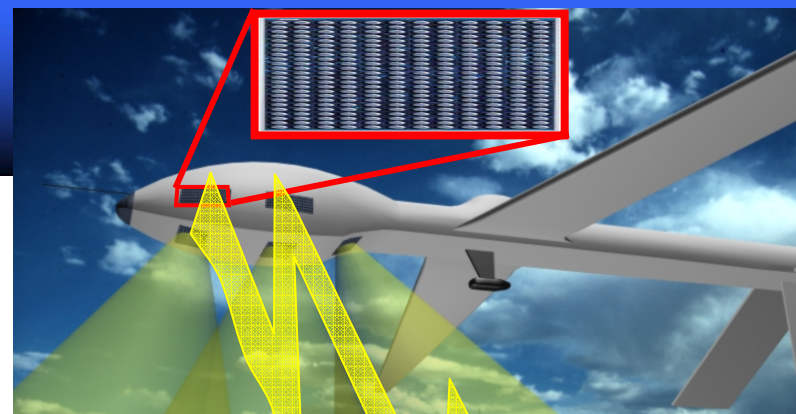
Sensor Technology: Too much of a good thing?



Compress: Compute hard, throw away 80% of the Megapixels for Downlink through tiny pipe



Reality:
High Resolution Image :
Big Optics, Megapixel Data



Vision:
Compressed Camera *directly measures*
Compressed representation for downlink

Uncomfortable Observation:

Compressibility means most of the megapixels are useless

BUT

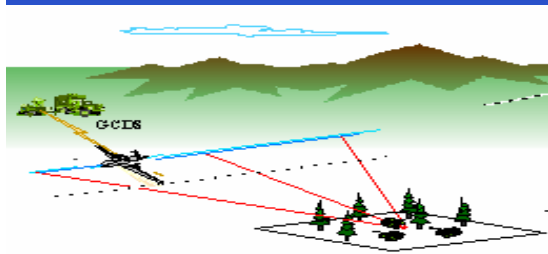
The gold is hidden in a broad spatial bandwidth, need smarts to find it!

Shannon: must measure the megapixels then sieve the hidden gold

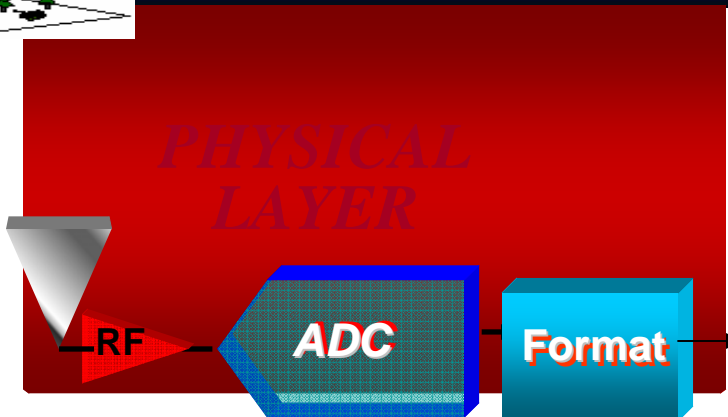
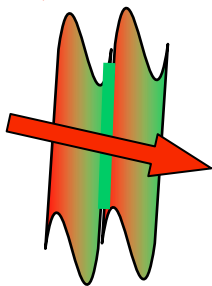
Perhaps Shannon had good reason to look worried?



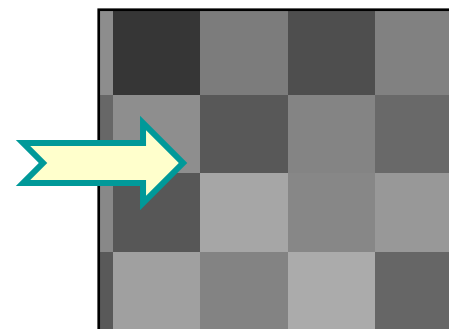
Sensor Trend: Measure/digitize everything possible Mega pixels = goodness!



Physical Field
(continuum)

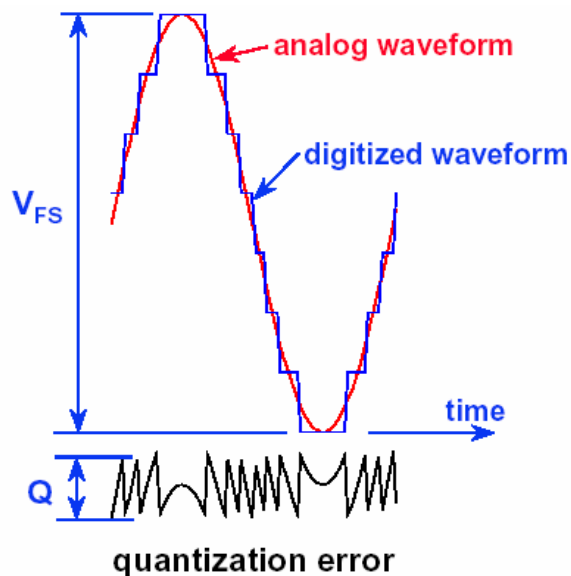


Digital Representation
(Formatted)

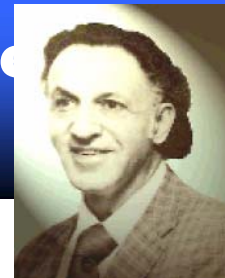


Raw sensor data:

*Huge Volume
low information/sample
Curse of Dimensionality*

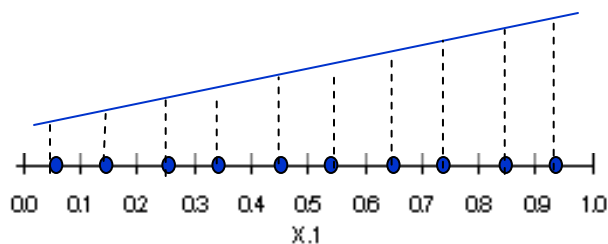


Bellman and the Curse of Dimensionality

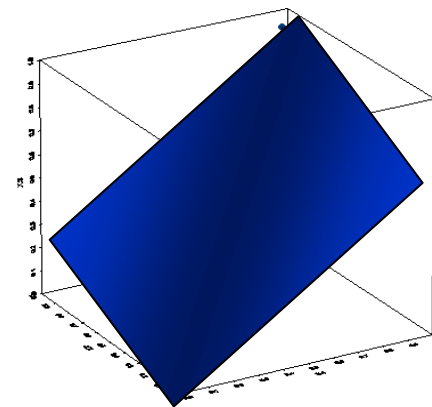


- Concept named by Richard Bellman

- Extraordinary growth in difficulty of doing business in higher dimensional spaces
 - Optimization, Approximation
 - Eg: “reasonable” function of d variables x_1, x_2, \dots, x_d $0 < x_i < 1$
 - $O(1/\epsilon)^d$ samples to obtain $O(\epsilon)$ accuracy



10 sample points are plenty in 1-d



Would need 100 sample points in the 2-d domain

- Statistics Version: Requirement on priors/training to do meaningful inference

High D, Ain't always the Place to Be

High dimensions are strange

The odd geometry of high d space

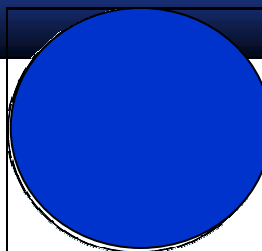
Wasteful packaging: The box for the high-d golf ball is essentially empty!

Volume is in the corners



volume of the hypercube is $(2r)^d$.

volume of the sphere is $\frac{2r^d \pi^{d/2}}{d \Gamma(d/2)}$.



$$\frac{\pi^{d/2}}{d 2^{d-1} \Gamma(d/2)} \rightarrow 0$$

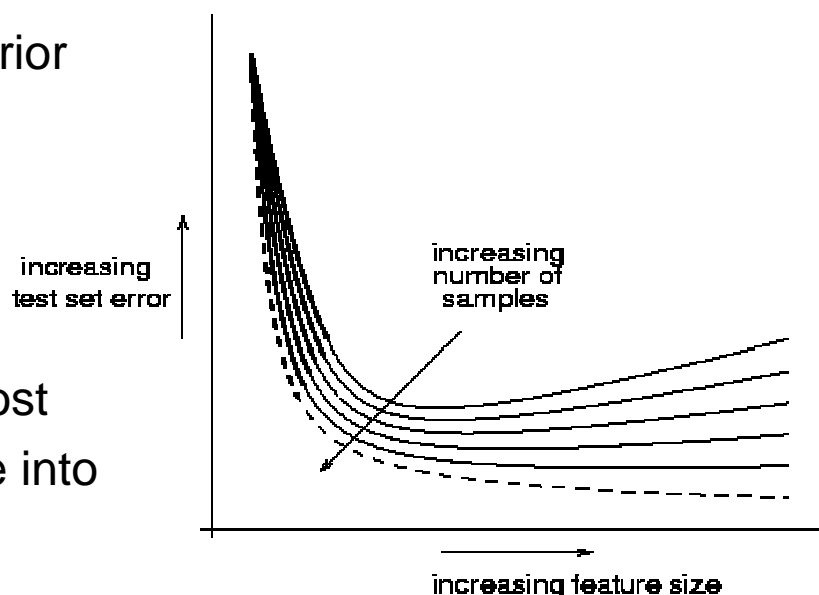
Most of uniform samples
are in the corners

Curses

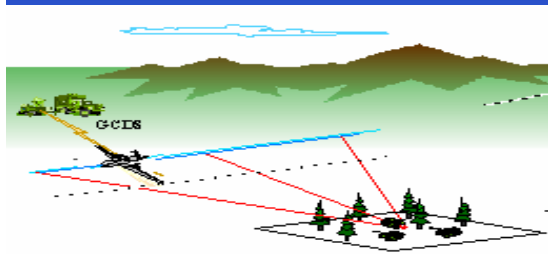
Dimensionality	Required Sample Size
1	4
2	19
5	786
7	10,700
10	842,000

Required samples to estimate
Gaussian density accurately at center

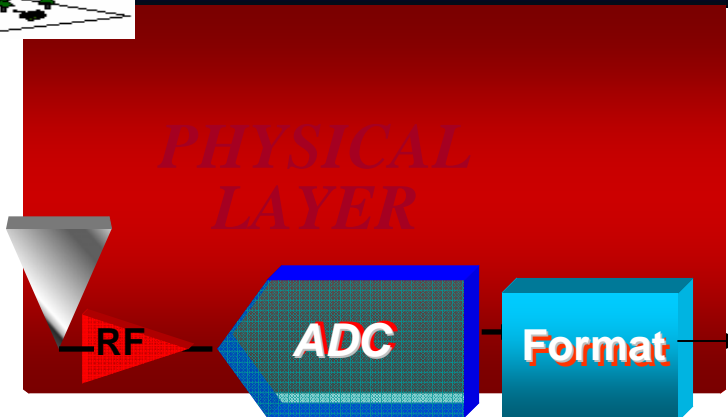
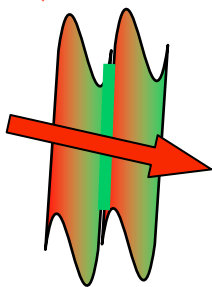
- Makes inference unreliable w/o more prior or training data
- Tends to make computation intractable
- If important structure is really low-d, most of the high-d features mainly drive noise into the application



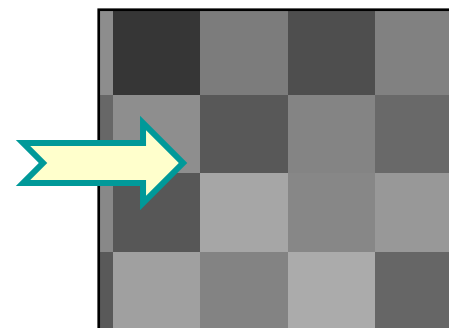
Bayes error



Physical Field
(continuum)

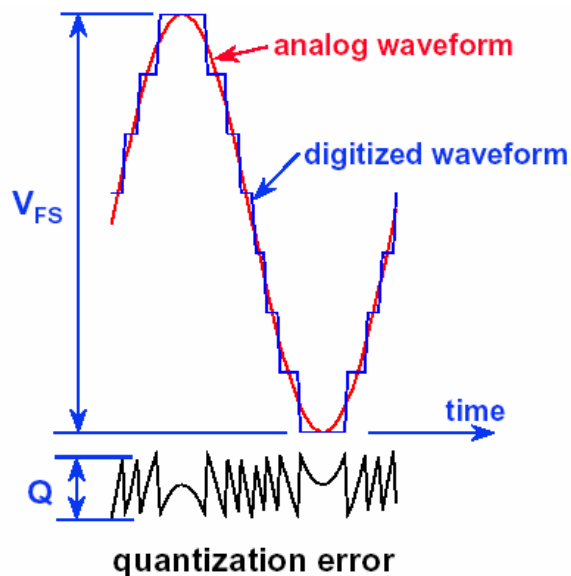


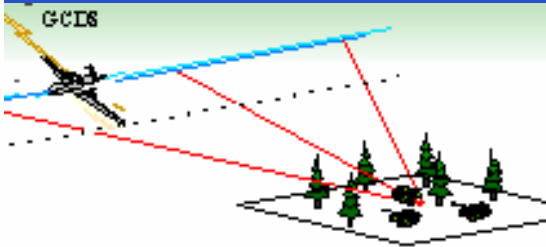
Digital Representation
(Formatted)



Raw sensor data:

*Huge Volume
low information/sample
Curse of Dimensionality*



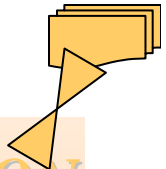
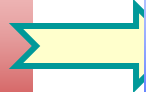
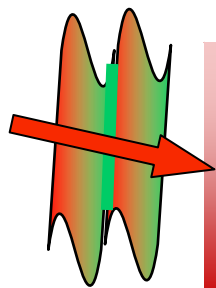


Physical Field
(continuum)

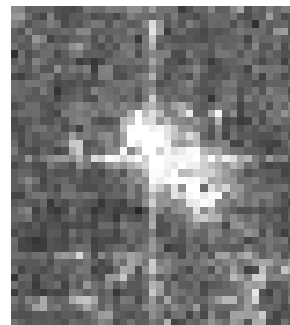
Digital Representation
(finite precision finite dimensional)

Transformed
Digital Representation

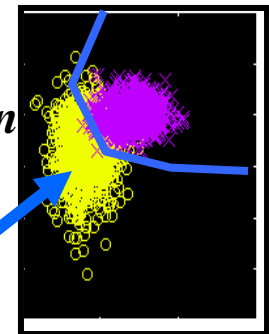
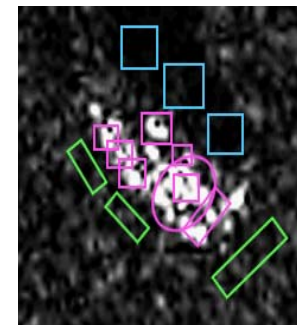
Symbolic Output



Raw sensor data:
low information/dimension

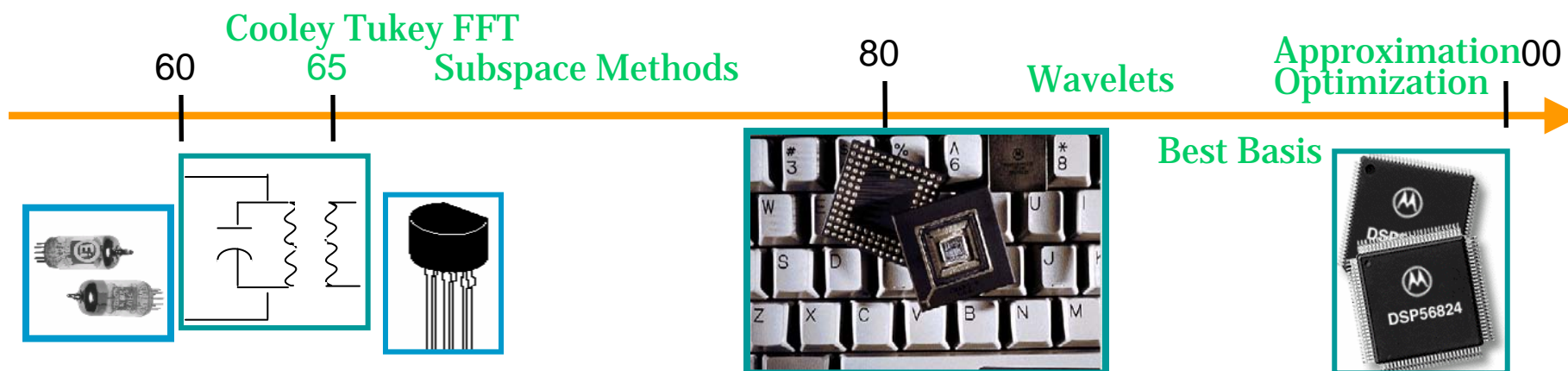
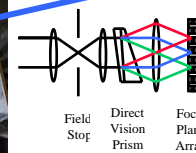
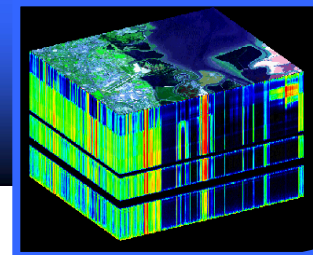
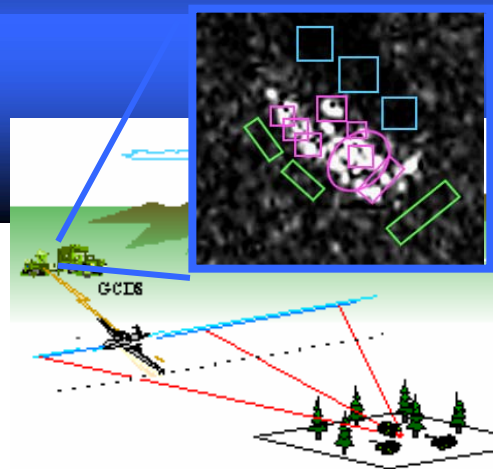
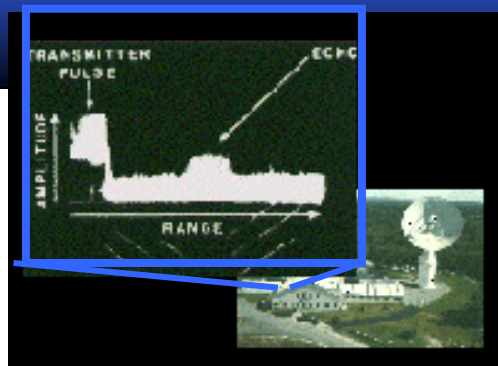


Enhanced feature data:
more information/dimension



Sensor: Work hard in the physical layer to measure everything at the highest fidelity

Processor: Work hard to throw away most of that data to get at the “good stuff”



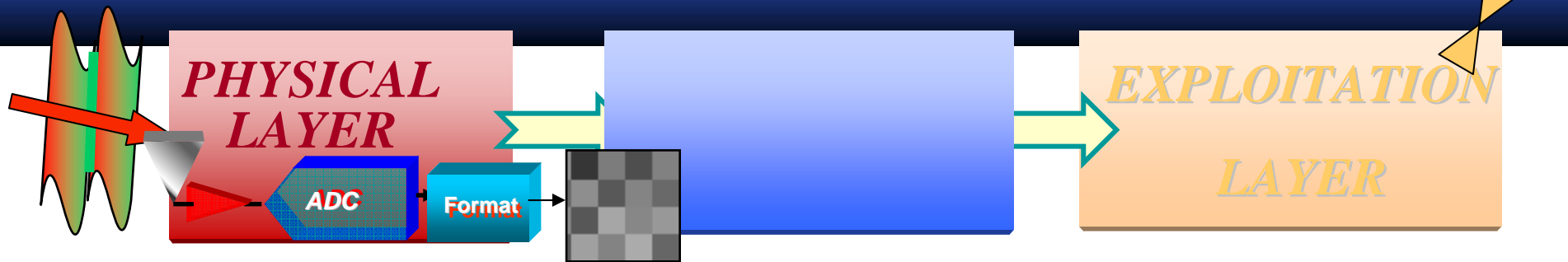
- *Increased complexity, dimensionality of sensing and exploitation tasks*
- *Decoupling of Digital and Analog subsystems, independently designed and optimized*
- *Steady Moore's Law growth of digital throughput (integration growth path)*
- *Equal impact of fast algorithms (representations), but sporadic, unpredictable*

Physical Field
(continuum)

Digital Representation
(finite precision finite dimensional)

Transformed
Digital Representation

Symbolic Output



Keep trying to measure/digitize everything at highest possible resolution?

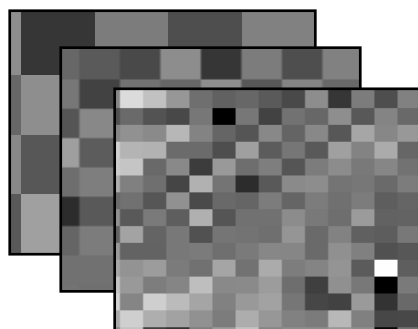
- A No-brainer
- No-brainer! Maybe not always true!
 - Slow, expensive growth rate of resolution in A/D's
 - Curse of dimensionality: shouldn't do it even if you could!
- What might be better?
 - Dedicate Front end resources to measuring more of the “good stuff”
 - Sensing “features” at front end could reduce load on A/D and subsequent digital processing and exploitation.
 - If done properly, overall system performance can be improved even with the lower requirements on subsequent digital processing
- **Problem: the “smarts” for finding the “good stuff” are behind the ADC!**

Today's Sensor Systems are typically feed-forward networks for transforming information in specialized stages.

Dimension reduction is mainly applied at the back of the bus



*Raw sensor data:
low information/dimension*

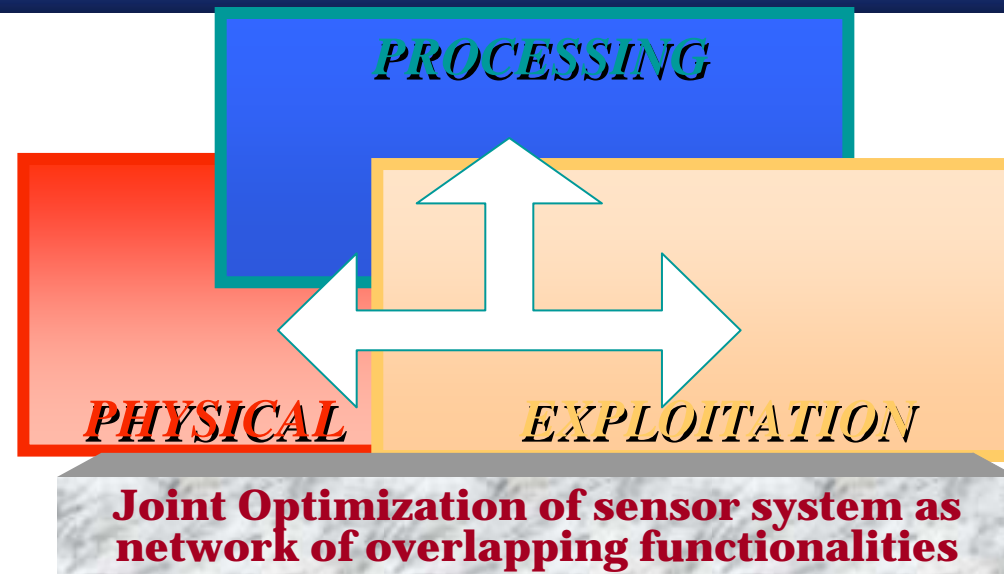


*DSP produces data reps with
higher information/dimension*



*Fast approximate
solution of a high
dimensional non
linear optimization
problem*

Meeting the Challenges: *Integrated Optimization of the System*



i

*Components will have overlapping and dynamically reconfigurable roles and full network connectivity. (**LOAD BALANCING**)*

*Make “customized” measurements at physical layer under real-time feedback control from the exploitation and processing system. **“20 Questions”***

*Manage/Prioritize **data** stream to affordable levels without discarding needed **information***